

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS





IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

METHOD FOR RECOGNIZING CHARACTER AND DEVICE THEREFOR

Patent number: JP6068301
Publication date: 1994-03-11
Inventor: WANG SHIN-YWAN; VAEZI MEHRZAD R; SHERRICK CHRISTOPHER A
Applicant: CANON INC.; CANON INF SYST INC
Classification:
 - international: G06K9/20; G06K9/32; G06K9/34
 - european:
Application number: JP19930121883 19930426
Priority number(s):

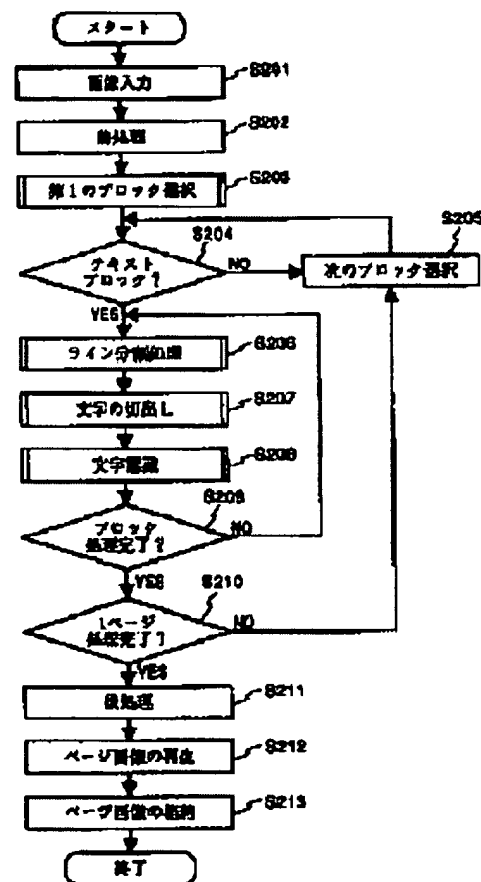
Also published as:

 EP0567344 (A2)
 US5680479 (A1)
 EP0567344 (A3)
 EP0567344 (B1)

Abstract of JP6068301

PURPOSE: To provide a character recognition method and device for recognizing a character on a document with a high speed and exactness, and preparing a text file.

CONSTITUTION: Pixel picture data of pixel units are inputted to a device (S201), the picture quality of the picture data is improved by a filter in order to improve deteriorated characters or pictures (S202), the hierarchical tree structure of the pictures is prepared (S203) so that text information/graphics information/ line information/picture information of information in a block can be identified, and each part in the picture can be reproduced in a proper sequence like a post-processing (S212), and whether or not the selected block is a text block is checked (S204). Then, a line dividing processing is operated to the text block (S206), each character in the line is segmented from another character in the line (S207), whether or not the processing to the text block and the processing to one page is ended is checked (S209 and 210), and a correction processing from an overall point of view such as context check or spelling check is operated (S211).



Data supplied from the esp@cenet database - Worldwide

T S1/5/1

1/5/1

DIALOG(R)File 347:JAPIO

(c) 2004 JPO & JAPIO. All rts. reserv.

04424401

METHOD AND DEVICE FOR RECOGNIZING CHARACTER

PUB. NO.: 06-068301 [JP 6068301 A]

PUBLISHED: March 11, 1994 (19940311)

INVENTOR(s): SHIN YAN WANGU

MEZAADOU AARU BAEZUII

KURISUTOFUAA EE SHIERITSUKU

APPLICANT(s): CANON INC [000100] (A Japanese Company or Corporation), JP
(Japan)

CANON INF SYST INC [000000] (A Non-Japanese Company or
Corporation), US (United States of America)

APPL. NO.: 05-121883 [JP 93121883]

FILED: April 26, 1993 (19930426)

PRIORITY: 7-873,012 [US 873012-1992], US (United States of America),
April 24, 1992 (19920424)

INTL CLASS: [5] G06K-009/20; G06K-009/32; G06K-009/34

JAPIO CLASS: 45.3 (INFORMATION PROCESSING -- Input Output Units); 44.7
(COMMUNICATION -- Facsimile)

JAPIO KEYWORD: R002 (LASERS); R107 (INFORMATION PROCESSING -- OCR & OMR
Optical Readers); R108 (INFORMATION PROCESSING -- Speech
Recognition & Synthesis); R131 (INFORMATION PROCESSING --
Microcomputers & Microprocessers); R139 (INFORMATION
PROCESSING -- Word Processors)

?

(19) 日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11) 特許出願公開番号

特開平6-68301

(43) 公開日 平成 6 年 (1994) 3 月 11 日

(51) Int.Cl.⁵

G 0 6 K 9/20
9/32
9/34

識別記号

3 4 0 L

庁内整理番号

F 1

技術表示箇所

審査請求 未請求 請求項の数286(全 420 頁)

(21) 出願番号 特願平5-121883

(22) 出願日 平成 5 年 (1993) 4 月 26 日

(31) 優先権主張番号 8 7 3 0 1 2

(32) 優先日 1992 年 4 月 24 日

(33) 優先権主張国 米国 (U S)

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子 3 丁目 30 番 2 号

(71) 出願人 592208172

キヤノン インフォメーション システム
ズ インク.

Canon Information S
ystems, Inc.

アメリカ合衆国 カリフォルニア州

92626, コスタ メサ, ブルマン スト
リート 3188

(74) 代理人 弁理士 大塚 康徳 (外 1 名)

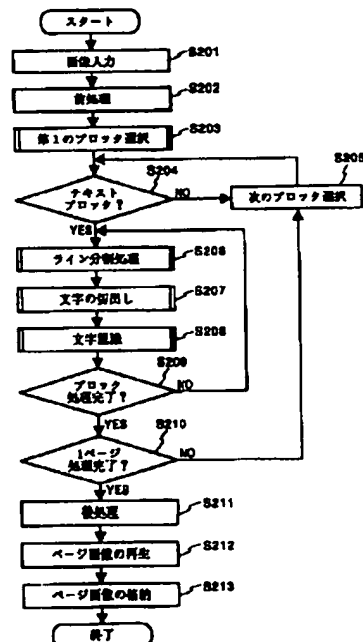
最終頁に続く

(54) 【発明の名称】 文字認識方法及び装置

(57) 【要約】 (修正有)

【目的】 高速かつ正確に文書上の文字を認識してテキストファイルを作成する文字認識方法及び装置の提供。

【構成】 画素単位の画素画像データを装置に入力し (S 2 0 1) 、 フィルターで画像データの画質を向上して劣化した文字や画像を改善し (S 2 0 2) 、 ブロック中情報のテキスト情報/グラフィックス情報/線画情報/画像情報などの識別と画像中の各部分を後処理 (S 2 1 2) のような適切な順序で再生できるように画像の階層的木構造を生成し (S 2 0 3) 、 選択されたブロックがテキストブロックか否かを調べ (S 2 0 4) 、 テキストブロックに対してライン分割処理を行い (S 2 0 6) 、 ライン中の各文字をライン中の他の文字から切り出し (S 2 0 7) 、 テキストブロックに対する処理と 1 ページ分に対する処理が終了したかを調べ (S 2 0 9 , 2 1 0) 、 文脈チェックやスペルチェックなどの全体的な見地からの修正処理を行う (S 2 1 1) 。



1

【特許請求の範囲】

【請求項1】 画素画像データから画素ブロックを選別する方法であって、

画素データ内で連結要素の輪郭を追跡する行程と、
輪郭が追跡された各連結要素を囲んで矩形を形成する行程と、

サイズと他の矩形への近接関係とに基づいて矩形を選択的に横方向に連結して、ラインを形成する第1の連結行程と、

サイズと他のラインへの近接関係とに基づいてラインを選択的に縦方向に連結して、ブロックを形成する第2の連結行程とを備える方法。

【請求項2】 更に、画素画像データを入力する行程を備え、

該入力行程は、画素画像データが2値画素画像データでない場合に、画素画像データを2値画素画像データに変換する行程を含む請求項1記載の方法。

【請求項3】 前記矩形形成行程では、各連結要素を囲む最小の矩形が形成される請求項1記載の方法。

【請求項4】 更に、前記矩形形成行程で形成された対応する矩形の位置に基づいて、前記輪郭追跡行程で追跡された連結要素の階層的木構造を形成する行程を備える請求項1記載の方法。

【請求項5】 更に、前記矩形形成行程で形成された矩形をテキスト要素と非テキスト要素とに分類する行程を備える請求項4記載の方法。

【請求項6】 更に、テキスト要素のブロック内の文字画像を認識する行程を備える請求項5記載の方法。

【請求項7】 前記分類行程は、要素の高さの所定のしきい値に対応して実行される請求項5記載の方法。

【請求項8】 前記分類行程は、前記矩形形成行程で形成された矩形の高さに関連する統計情報に対応して実行される請求項6記載の方法。

【請求項9】 前記第1の連結行程と第2の連結行程とは、非テキスト要素に対しては実行されない請求項6記載の方法。

【請求項10】 更に、非テキスト要素に対し白輪郭を得る行程を備える請求項6記載の方法。

【請求項11】 非テキスト要素は白輪郭の数に対応して表要素と見なされる請求項10記載の方法。

【請求項12】 更に、白輪郭の充填率を算出する行程を備える請求項10記載の方法。

【請求項13】 非テキスト要素は、充填率が高い場合には画像データと見なされない請求項12記載の方法。

【請求項14】 更に、非格子状に配置された白輪郭を再接続する行程を備える請求項12記載の方法。

【請求項15】 非テキスト要素は、再接続率が高くない場合に表と見なされる請求項14記載の方法。

【請求項16】 白輪郭は4方向で算出される請求項10記載の方法。

2

【請求項17】 連結要素の輪郭は、少なくとも8方向に追跡される請求項1記載の方法。

【請求項18】 前記輪郭追跡行程では、連結要素の外部のみで連結要素の輪郭が追跡される請求項1記載の方法。

【請求項19】 更に、画素画像データ内でギャップを検出する行程を備え、前記第1の連結行程では、ギャップが矩形を隔てている場合には矩形がラインに連結されない請求項1記載の方法。

【請求項20】 矩形の間を縦に延びるギャップに対応して段を検出する請求項19記載の方法。

【請求項21】 前記第2の連結行程は、前記第1の連結行程で連結されたテキストデータのライン間に非テキストの境界を決定する行程を含み、前記第2の連結行程は、間に非テキストの境界がある場合にラインを縦方向にブロックとして連結しない請求項1記載の方法。

【請求項22】 更に、前記輪郭追跡行程の前に画素画像データを圧縮する行程を備える請求項1記載の方法。

【請求項23】 画素画像データ内の文字を認識する方法であって、

画素画像データ内で連結要素の輪郭を追跡する行程と、
追跡された連結要素がテキスト要素を含むか非テキスト要素を含むかを定める行程と、
テキスト要素を横方向に選択的に連結してテキストラインを形成する行程と、
テキストラインを選択的に縦方向に連結してテキストブロックを形成する行程とを含み、
画素画像データから画素ブロックを選別する行程と、
テキストブロックを画素画像データのラインに分割する行程と、

前記分割行程で分割されたラインから文字を切り出す行程と、

前記切り出行程で切り出された文字を認識する行程と、

前記認識行程で認識された文字を、前記ブロック選別行程で確定された順に対応して格納する行程とを備える方法。

【請求項24】 更に、テキスト要素のブロック内の文字画像を認識する行程を備える請求項23記載の方法。

【請求項25】 更に、画素画像データが前処理される前処理行程を備える請求項23記載の方法。

【請求項26】 前記前処理行程は画像圧縮行程を含む請求項25記載の方法。

【請求項27】 前記前処理行程は画素画像データの画質を向上するフィルタ行程を含む請求項25記載の方法。

【請求項28】 更に、前記認識行程で認識された文字を後処理する行程を備える請求項23記載の方法。

【請求項29】 前記後処理は文脈チェックを含む請求項28記載の方法。

【請求項30】 更に、画素画像データを入力する行程を備え、

3

該入力行程は、画素画像データが2値画素画像データでない場合に、画素画像データを2値画素画像データに変換する行程を含む請求項2記載の方法。

【請求項31】 前記ブロック選別行程は、前記輪郭追跡行程で定められた連結画素に基づいて階層の木構造を形成する行程を含む請求項2記載の方法。

【請求項32】 更に、前記輪郭追跡行程で追跡された連結要素を囲む矩形を形成する行程を備える請求項31記載の方法。

【請求項33】 前記階層の木構造は、前記矩形形成行程で形成された矩形の位置に基づいて形成される請求項32記載の方法。

【請求項34】 更に、第1及び第2の連結行程に基づいて、階層の木構造を更新する行程を備える請求項31記載の方法。

【請求項35】 更に、連結要素をテキスト要素又は非テキスト要素に分類する行程を備える請求項31記載の方法。

【請求項36】 前記第1及び第2の連結行程は、テキスト要素を連結するが非テキスト要素は連結しない請求項35記載の方法。

【請求項37】 更に、非テキスト要素の内部で白輪郭を追跡する行程を備える請求項35記載の方法。

【請求項38】 更に、非テキスト要素に含まれる白輪郭の数に基づいて、非テキスト要素を表と見なす行程を備える請求項37記載の方法。

【請求項39】 更に、非格子状に配置された白輪郭を再接続する行程を備える請求項37記載の方法。

【請求項40】 前記非テキスト要素は、前記再接続行程での再接続に基づいて表と見なされる請求項39記載の方法。

【請求項41】 更に、非テキスト要素内の白輪郭の充填率を算出する行程を備える請求項35記載の方法。

【請求項42】 非テキスト要素は、充填率が高い場合に表と見なされない請求項41記載の方法。

【請求項43】 画素画像データ内で文字のテキストファイルを形成する方法であって、画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する行程を含む画素画像データの入力行程と、画素画像データ内で連結要素の輪郭を追跡する行程と、追跡された連結要素のサイズに基づいて連結要素がテキスト要素か非テキスト要素かを定める行程と、隣接するテキスト要素の近接関係に基づいてテキスト要素を選択的に横方向に連結してテキストラインを形成する行程と、隣接するテキストラインの近接関係とテキスト要素間の非テキスト要素の位置とに基づいてテキストラインを選択的に縦方向に連結してテキストブロックを形成する行程とを含み、画素画像データのブロックを選別する行程と、

4

前記輪郭追跡で追跡された画素要素に基づいて、階層の木構造を形成する行程と、

テキストブロックを画素画像データのラインに分割する行程と、

前記分割行程で分割されたラインから文字を切り出す行程と、

前記切り出行程で切り出された文字を認識し、前記認識に基づいて文字コードを割り当てる行程と、

前記階層の木構造で確定された順に対応して文字コードをコンピュータのテキストファイルに格納する行程とを備える方法。

【請求項44】 更に、画素画像データが前処理される前処理行程を備える請求項43記載の方法。

【請求項45】 前記前処理行程は画像圧縮行程を含む請求項44記載の方法。

【請求項46】 前記前処理行程は画素画像データの画質を向上する行程を含む請求項45記載の方法。

【請求項47】 更に、文脈チェックを含み、前記認識行程で認識された文字を後処理する行程を備える請求項43記載の方法。

【請求項48】 更に、非テキスト要素の内部で白輪郭を追跡する行程を備える請求項43記載の方法。

【請求項49】 更に、非テキストに識別子を付与する行程を更に備える請求項48記載の方法。

【請求項50】 更に、非テキスト要素に含まれる白輪郭の数に基づいて表識別子を付与する行程を備える請求項49記載の方法。

【請求項51】 更に、非格子状配置の白輪郭を再接続する行程を備え、非テキスト要素には前記再接続行程での再接続率に基づいて表の識別子が付与される請求項49記載の方法。

【請求項52】 更に、非テキスト要素内の白輪郭の充填率を算出する行程を備え、充填率が低い場合に非テキスト要素に表の識別子を付与する請求項49記載の方法。

【請求項53】 画素画像データから画素ブロックを選別する装置であって、画素データ内で連結要素の輪郭を追跡する輪郭追跡手段と、

前記輪郭追跡手段によって追跡された各連結要素を囲んで矩形を形成する矩形形成手段と、

サイズと他の矩形への近接関係とに基づいて矩形を選択的に横方向に連結して、ラインを形成する第1連結手段と、

サイズと他のラインへの近接関係に基づいてラインを選択的に縦方向に連結して、ブロックを形成する第2連結手段とを備える装置。

【請求項54】 更に、画素画像データを入力する入力手段を備え、

該入力手段は、画素画像データが2値画素画像データで

ない場合に、画素画像データを2値画素画像データに変換する変換手段を含む請求項53記載の装置。

【請求項55】 前記矩形形成手段では、各連結要素を囲む最小の矩形が形成される請求項53記載の装置。

【請求項56】 更に、前記矩形形成手段で形成された対応する矩形の位置に基づいて、前記輪郭追跡手段で追跡された連結要素の階層の木構造を形成する第2の形成手段を備える請求項53記載の装置。

【請求項57】 更に、前記矩形形成手段で形成された矩形をテキスト要素と非テキスト要素とに分類する分類手段を備える請求項56記載の装置。

【請求項58】 更に、テキスト要素のブロック内の文字画像を認識する認識手段を備える請求項57記載の装置。

【請求項59】 前記分類手段は、要素の高さの所定のしきい値に対応して分類する請求項57記載の装置。

【請求項60】 前記分類手段は、前記矩形形成手段で形成された矩形の高さに関連する統計情報に対応して分類する請求項57記載の装置。

【請求項61】 前記第1連結手段と第2連結手段とは、非テキスト要素には使用されない請求項57記載の装置。

【請求項62】 更に、非テキスト要素に対して白輪郭を得る手段を備える請求項57記載の装置。

【請求項63】 非テキスト要素は白輪郭の数に対応して表の要素と見なされる請求項62記載の装置。

【請求項64】 更に、白輪郭の充填率を算出する算出手段を備える請求項62記載の装置。

【請求項65】 非テキスト要素は、充填率が高い場合には画像データと見なされない請求項64記載の装置。

【請求項66】 更に、非格子状に配置された白輪郭を再接続する再接続手段を備える請求項64記載の装置。

【請求項67】 非テキスト要素は、再接続率が高くない場合に表と見なされる請求項66記載の装置。

【請求項68】 白輪郭は4方向で算出される請求項62記載の装置。

【請求項69】 連結要素の輪郭は、少なくとも8方向に追跡される請求項53記載の装置。

【請求項70】 前記輪郭追跡手段では、連結要素の外部のみで連結要素の輪郭が追跡される請求項53記載の装置。

【請求項71】 更に、画素画像データ内でギャップを検出する検出手段を備え、前記第1連結手段では、ギャップが矩形を隔てている場合には矩形がラインに連結されない請求項53記載の装置。

【請求項72】 矩形の間を縦に延びるギャップに対応して段が検出される請求項71記載の装置。

【請求項73】 前記第2連結手段は、前記第1連結手段で連結されたテキストデータのライン間に非テキストの境界を決定する決定手段を含み、前記第2連結手段

は、間に非テキストの境界がある場合にラインを縦方向にブロックとして連結しない請求項53記載の装置。

【請求項74】 更に、前記輪郭追跡手段の前に、画素画像データを圧縮する圧縮手段を備える請求項53記載の装置。

【請求項75】 画素画像データ内の文字を認識する装置であって、

画素画像データ内で連結要素の輪郭を追跡する行程と、定められた連結要素がテキスト要素を含むか非テキスト要素を含むかを定める行程と、テキスト要素を横方向に選択的に連結してテキストラインを形成する行程と、テキストラインを選択的に縦方向に連結してテキストブロックを形成する行程とを含み、画素画像データから画素ブロックを選別するブロック選別手段と、

テキストブロックを画素画像データのラインに分割する分割手段と、

前記分割手段で分割されたラインから文字を切り出す切出手段と、

前記切出手段で切り出された文字を認識する認識手段と、

前記認識手段で認識された文字を、前記ブロック選別手段で確定された順に対応して格納する格納手段とを備える装置。

【請求項76】 更に、テキスト要素のブロック内の文字画像を認識する第2の認識手段を備える請求項75記載の装置。

【請求項77】 更に、画素画像データが前処理される前処理手段を備える請求項75記載の装置。

【請求項78】 前記前処理手段は画像圧縮手段を含む請求項77記載の装置。

【請求項79】 前記前処理手段は画素画像データの画質を向上するフィルタ手段を含む請求項77記載の装置。

【請求項80】 更に、前記認識手段で認識された文字を後処理する後処理手段を備える請求項75記載の装置。

【請求項81】 前記後処理手段は文脈チェック手段を含む請求項80記載の装置。

【請求項82】 更に、画素画像データを入力する入力手段を備え、

該入力手段は、画素画像データが2値画素画像データでない場合に、画素画像データを2値画素画像データに変換する変換手段を含む請求項75記載の装置。

【請求項83】 前記ブロック選別手段は、前記輪郭追跡手段で追跡された連結要素に基づいて階層の木構造を形成する手段を含む請求項75記載の装置。

【請求項84】 更に、前記輪郭追跡手段で定められた連結要素を囲む矩形を形成する形成手段を備える請求項83記載の装置。

【請求項85】 前記階層の木構造は、前記矩形形成手

段で形成された矩形の位置に基づいて形成される請求項84記載の装置。

【請求項86】 更に、第1及び第2連結手段に基づいて、階層の木構造を更新する更新手段を備える請求項64記載の装置。

【請求項87】 更に、連結要素をテキスト要素又は非テキスト要素に分類する分類手段を備える請求項83記載の装置。

【請求項88】 前記第1及び第2連結手段は、テキスト要素を連結するが非テキスト要素は連結しない請求項10 87記載の装置。

【請求項89】 更に、非テキスト要素の内部で白輪郭を追跡する手段を備える請求項87記載の装置。

【請求項90】 更に、非テキスト要素に含まれる白輪郭の数に基づいて表と見なす手段を備える請求項89記載の装置。

【請求項91】 更に、非格子状に配置された白輪郭を再接続する再接続手段を備える請求項89記載の装置。

【請求項92】 前記非テキスト要素は、前記再接続手段での再接続に基づいて表と見なされる請求項91記載 20 の装置。

【請求項93】 更に、非テキスト要素内の白輪郭の充填率を算出する算出手段を備える請求項87記載の装置。

【請求項94】 非テキスト要素は、充填率が高い場合に表と見なされない請求項93記載の装置。

【請求項95】 画素画像データ内で文字のテキストファイル形成する装置であって、

画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する変換 30 手段を含む画素画像データの入力手段と、

画素画像データ内で連結要素の輪郭を追跡する輪郭追跡手段と、取り出された連結要素のサイズに基づいて連結要素がテキスト要素か非テキスト要素かを決定する決定手段と、隣接するテキスト要素の近接関係に基づいてテキスト要素を選択的に横方向に連結してテキストラインを形成する第1連結手段と、隣接するテキストラインの近接関係とテキスト要素間の非テキスト要素の位置とに基づいてテキストラインを選択的に縦方向に連結してテキストブロックを形成する第2連結手段とを含み、画素画像 40 データのブロックを選別するブロック選別手段と、

前記輪郭追跡で追跡された画素要素に基づいて、階層の木構造を形成する形成手段と、

テキストブロックを画素画像データのラインに分割する分割手段と、

前記分割手段で分割されたラインから文字を切り出す切出手段と、

前記切出手段で切り出された文字を認識し、前記認識に基づいて文字コードを割り当てる認識手段と、

前記階層の木構造で確定された順に対応して文字コード 50

をコンピュータのテキストファイルに格納する格納手段とを備える装置。

【請求項96】 更に、画素画像データが前処理される前処理手段を備える請求項95記載の装置。

【請求項97】 前記前処理手段は画像圧縮手段を含む請求項96記載の装置。

【請求項98】 前記前処理手段は画素画像データの画質向上手段を含む請求項97記載の装置。

【請求項99】 更に、文脈チェック手段を含み、前記認識手段で認識された文字を後処理する後処理手段を備える請求項95記載の装置。

【請求項100】 更に、非テキスト要素の内部で白輪郭を追跡する手段を備える請求項95記載の装置。

【請求項101】 更に、非テキストに識別子を付与する付与手段を更に備える請求項100記載の装置。

【請求項102】 更に、非テキスト要素に含まれる白輪郭の数に基づいて、非テキスト要素に表の識別子が付与される請求項101記載の装置。

【請求項103】 更に、非格子状配置の白輪郭を再接続する再接続手段を備え、非テキスト要素には前記再接続手段での再接続率に基づいて表の識別子が付与される請求項101記載の装置。

【請求項104】 更に、非テキスト要素内の白輪郭の充填率を算出する算出手段を備え、充填率が高くない場合に非テキスト要素に表識別子を付与する請求項101記載の装置。

【請求項105】 画素画像データでラインの位置を割り出す方法であって、

画像データの横方向に伸びた部分の全域で画像密度の水平投影を得る行程と、

前記水平投影に基づいて、画像密度が投影された部分の幅を減少するか否かを決定する行程と、

前記水平投影に基づいて、ラインの位置を割り出す行程を備える方法。

【請求項106】 前記決定行程は、実質的に空白の部分で水平投影を調べる行程と、実質的に空白の部分が見付からない場合に幅を減少する行程とを備える請求項105記載の方法。

【請求項107】 前記調べる行程は、間隔の小さい非空白の部分を選択する行程を含む請求項106記載の方法。

【請求項108】 前記水平投影を得る行程と決定する行程とは、順に繰り返して実行される請求項105記載の方法。

【請求項109】 前記割り出行程では、水平投影の実質的に空白の領域間の画像データのライン位置を、ラインとして割り出す請求項105記載の方法。

【請求項110】 更に、画素画像データを前処理する前処理行程を備える請求項105記載の方法。

【請求項111】 段に分割可能な画素画像データ内で

ラインの位置を割り出す方法であって、画像データの少なくとも1つの横方向に伸びる段の全域で、画像密度の水平投影を得る行程と、前記行程で得られた水平投影に基づいて、画像密度が投影される段の数を増加するか否かを決定する行程と、前記水平投影に基づいてデジタル画像データ内でラインの位置を割り出す行程とを備える方法。

【請求項112】 前記決定行程は、実質的に空白の部分で水平投影を調べる行程と、実質的に空白の部分が見付からない場合に、段の数を増加する行程とを備える請求項111記載の方法。

【請求項113】 前記調べる行程は間隔の小さい非空白の部分で連結する行程を含む請求項112記載の方法。

【請求項114】 前記水平投影を得る行程と決定する行程とは、順に繰り返して実行される請求項111記載の方法。

【請求項115】 前記割出行程では、水平投影の実質的に空白の領域間の画像データのライン位置を、ラインとして割り出す請求項111記載の方法。

【請求項116】 更に、画像データの隣接する段のラインが重なるか否かを決定する行程を備える請求項111記載の方法。

【請求項117】 前記ラインが重なるかを決定する行程は、スケール不変の決定から成る請求項116記載の方法。

【請求項118】 更に、画素画像データを前処理する前処理行程を備え、前記前処理行程は画像縮小行程を含む請求項111記載の方法。

【請求項119】 前記画像縮小は水平方向と垂直方向とで異なる請求項118記載の方法。

【請求項120】 画素画像データのラインセグメントから文字画像を切り出す方法であって、ラインセグメントから非接触で重なっていない文字を切り出す第1の切出行程と、ラインセグメントから接触文字を切り出す第2の切出行程とを備える方法。

【請求項121】 更に、高解像度の画像セグメントを低解像度に圧縮する画像データのイメージセグメントの前処理行程を備える請求項120記載の方法。

【請求項122】 垂直方向と水平方向とで異なる圧縮率が使われる請求項121記載の方法。

【請求項123】 更に、前記第1の切出行程で切り出された文字を認識処理する行程を備え、前記第2の切出行程は前記認識行程で認識されなかった文字に対してのみ実行される請求項120記載の方法。

【請求項124】 更に、非接触で重なっている文字データをラインセグメントから切り出す中間の切出行程を備える請求項123記載の方法。

【請求項125】 前記中間切出行程は、非接触で重な

っている文字を画素データの輪郭の追跡により切り出す行程を含む請求項124記載の方法。

【請求項126】 前記中間切出行程は、間違っ切り出された文字を再接続する行程を含む請求項124記載の方法。

【請求項127】 更に、前記認識行程で認識されなかった文字を再接続する行程を備える請求項123記載の方法。

【請求項128】 前記第1の切出行程は、前記ラインセグメントを非空白の画素データが見付かるまでとびとびに処理する行程を含む請求項120記載の方法。

【請求項129】 更に、前記とびとびの処理行程で非空白の画素が見付かった場合に、厳密に前後をサーチする行程を備える請求項128記載の方法。

【請求項130】 前記第2の切出行程は非垂直切り出しによって接触文字を切り出す行程を含む請求項120記載の方法。

【請求項131】 非垂直切り出しの角度は画素密度の垂直投影特性に対応して決められる請求項130記載の方法。

【請求項132】 更に、垂直投影特性に基づいて少なくとも1つの回転された投影特性を得る行程を備えることを特徴とする請求項131記載の方法。

【請求項133】 更に、前記垂直投影特性の角度に近接する角度で複数の回転された投影特性を得る行程を備える請求項132記載の方法。

【請求項134】 非垂直切り出しの角度は、前記回転投影特性を得る行程と垂直投影特性を得る行程とから得られた投影特性の最小値に基づいて得られる請求項133記載の方法。

【請求項135】 前記第2の切出行程は、前記ラインセグメント内の文字の特徴が既知かどうかに基づいて選択可能である請求項120記載の方法。

【請求項136】 前記第2の切出行程は、前記文字セット内の文字データのスペースの統計に基づいて切り出す行程を含む請求項135記載の方法。

【請求項137】 更に、小さな文字のセットを再接続する行程を備える請求項136記載の方法。

【請求項138】 文字画素データのラインセグメント内で接触文字を切り出す方法であって、ラインセグメント内で画素データの垂直投影特性を得る行程と、

垂直投影特性の最小値から隣接する最大値への方向から少なくとも1つの角度を算出する第1の算出行程と、

前記第1の算出行程で算出された角度に基づいて、少なくとも1つの回転された投影特性を算出する第2の算出行程と、

前記回転投影特性の最小値に対応する位置で前記回転投影特性の角度にラインセグメントを切る行程とを備える方法。

【請求項139】 前記第1の算出行程は、前記垂直投影特性を複数のしきい値と比較して前記最小値と隣接する最大値との位置を決定する行程を含む請求項138記載の方法。

【請求項140】 更に、最小値と隣接する最大値とが見付からない場合に、前記しきい値を増加させる行程を備える請求項139記載の方法。

【請求項141】 最小値は、少なくとも一方の側が第2しきい値より大きい最大値によって囲まれ、第1しきい値より小さな垂直投影特性上の位置として割り出される請求項139記載の方法。

【請求項142】 更に、各々が前記第2の算出行程で算出された回転投影特性に近接する角度で算出される複数の回転投影特性を算出する行程を備える請求項138記載の方法。

【請求項143】 更に、複数の回転投影特性と垂直投影特性との全体の最小値を割り出す行程を備え、前記切断行程は最小値の位置に対応する位置で最小値が得られた角度に対応する角度に切る請求項142記載の方法。

【請求項144】 画素画像データ内で文字を認識する方法であって、

画素画像データから画素ブロックを選別する行程と、少なくとも1つの段全体の画素密度の水平投影に基づいてテキストブロックを少なくとも1つの段に適応的に分配することにより、テキストブロックを画素データのラインに分割する行程と、

非接触で重なっていない文字がラインセグメントから切り出される第1の切出行程と、接触文字を切り出す少なくとも1つの追加の切出行程とを含み、前記分割行程で分割されたラインセグメントから文字を切り出す行程と、

前記文字切出行程で切り出された文字を認識する行程と、

前記認識行程で認識された文字を前記ブロック選別行程で確定された順に対応して格納する行程とを備える方法。

【請求項145】 更に、画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する行程を含む画素画像データを入力する行程を備える請求項144記載の方法。

【請求項146】 更に、画素画像データが前処理される前処理行程を備える請求項144記載の方法。

【請求項147】 前記前処理行程は、画像圧縮行程を含む請求項146記載の方法。

【請求項148】 前記前処理行程は、画素画像データの画質を向上するフィルタ行程を含む請求項146記載の方法。

【請求項149】 更に、前記認識行程で認識された文字を後処理する行程を備える請求項144記載の方法。

【請求項150】 前記後処理は文脈チェックを含む請求

項149記載の方法。

【請求項151】 前記第1の切出行程はラインセグメントをとびとびに処理する行程を含む請求項144記載の方法。

【請求項152】 前記追加の切出行程は、画素画像データの文字について特徴が既知か否かに対応して選択可能である請求項144記載の方法。

【請求項153】 前記追加の切出行程では既知の文字セットの統計に基づいて切り出しを行う請求項152記載の方法。

【請求項154】 前記追加の切出行程では画素密度の回転投影に対応して切り出しを行う請求項152記載の方法。

【請求項155】 画素密度の回転投影は画素密度の垂直投影に基づいて決められた角度で得られる請求項154記載の方法。

【請求項156】 更に、前記第1の切出行程と前記少なくとも1つの追加の切出行程の間に中間の切出行程を備える請求項154記載の方法。

【請求項157】 前記中間の切出行程はラインセグメント内の重なった画素画像の輪郭を追跡する行程を含む請求項156記載の方法。

【請求項158】 更に、文字が間違っって切り出された場合に、前記中間の切出行程で切り出された文字を再接続する行程を備える請求項157記載の方法。

【請求項159】 前記認識行程は、前記追加の切出行程の前に前記第1の切出行程で切り出された文字を認識し、前記追加の切出行程は前記認識行程で認識されない文字に対して行なわれる請求項158記載の方法。

【請求項160】 更に、文字が前記認識行程で認識されない場合に、前記少なくとも1つの追加の切出行程で切り出された文字を再接続する行程を備える請求項159記載の方法。

【請求項161】 画素画像データから文字のテキストファイルを形成する方法であって、

画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する行程を含む画素画像データの入力行程と、

画素画像データのブロックがテキストブロックか非テキストブロックかに対応して画素画像データのブロックを選別する行程と、

テキスト及び非テキストブロックの階層的木構造を得る行程と、

少なくとも1つの段の全域の画素密度の水平投影に基づいて、テキストブロックを前記少なくとも1つの段で適応的に分離することにより、テキストブロックを画素画像データのラインに分割する行程と、

非接触で重なっていない文字をラインセグメントから切り出す第1の切出行程と、接触文字を切り出す少なくとも1つの追加の切出行程とを含み、前記分割行程で分割

されたラインから文字を切り出す行程とを備える方法。

【請求項162】 更に、画素画像データが前処理される前処理行程を備える請求項161記載の方法。

【請求項163】 前記前処理行程は画像圧縮行程を含む請求項162記載の方法。

【請求項164】 前記前処理行程は画素画像データの画質を向上する行程を含む請求項163記載の方法。

【請求項165】 更に、文脈チェックを含み、前記認識行程で認識された文字を後処理する行程を備える請求項161記載の方法。

【請求項166】 前記第1の切出行程はラインセグメントをとびとびに処理する行程を含む請求項161記載の方法。

【請求項167】 前記追加の切出行程は、画素画像データの文字について特徴が既知か否かに対応して選択可能である請求項166記載の方法。

【請求項168】 前記追加の切出行程では既知の文字セットの統計に基づいて切り出しを行う請求項167記載の方法。

【請求項169】 前記追加の切出行程では画素密度の回転投影に対応して切り出しを行う請求項167記載の方法。

【請求項170】 画素密度の回転投影は画素密度の垂直投影に基づいて決められた角度で得られる請求項169記載の方法。

【請求項171】 更に、前記第1の切出行程と前記少なくとも1つの追加の切出行程の間に中間の切出行程を備える請求項161記載の方法。

【請求項172】 前記中間の切出行程はラインセグメント内の重なった画素画像の輪郭を追跡する行程を含む請求項171記載の方法。

【請求項173】 更に、文字が間違っって切り出された場合に、前記中間の切出行程で切り出された文字を再接続する行程を備える請求項172記載の方法。

【請求項174】 前記認識行程は、前記追加の切出行程の前に前記第1の切出行程で切り出された文字を認識し、前記追加の切出行程は前記認識行程で認識されない文字に対して行なわれる請求項173記載の方法。

【請求項175】 更に、文字が前記認識行程で認識されない場合に、前記少なくとも1つの追加の切出行程で切り出された文字を再接続する行程を備える請求項174記載の方法。

【請求項176】 画素画像データから文字のテキストファイルを形成する方法であって、画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する行程を含み、画素画像データを入力する行程と、

画素画像データ内で連結要素の輪郭を追跡する行程と、前記追跡行程の間に追跡された連結要素のサイズに基づいて追跡された連結要素がテキスト要素を含むか非テキ

スト要素を含むから決める行程と、隣接するテキスト要素の近接関係に基づいてテキスト要素を選択的に横方向に連結してテキストラインを形成する行程と、隣接するテキストラインの近接関係とテキストライン間の非テキスト要素の位置とに基づいてテキストラインを選択的に縦方向に連結する行程とを含み、画素画像データのブロックを選別する行程と、

前記輪郭追跡行程で追跡された連結要素に基づいて、階層の木構造を形成する行程と、

10 少なくとも1つの段の全域での画素密度の水平投影に基づいて、テキストブロックを前記少なくとも1つの段に適応的に分離することによって、テキストブロックを画素データのラインに分割する行程と、

非接触で重なっていない文字をラインセグメントから切り出す第1の切出行程と、接触文字を切り出す少なくとも1つの追加の切出行程とを含み、前記分割行程で分割されたラインから文字を切り出す行程と、

前記文字切出行程で切り出された文字を認識し、前記認識に基づいて文字コードを割り付ける行程と、

20 階層の木構造により確定された順に対応して、文字コードをコンピュータのテキストファイルに格納する行程とを備える方法。

【請求項177】 更に、画素画像データが前処理される前処理行程を備える請求項176記載の方法。

【請求項178】 前記前処理行程は画像圧縮行程を含む請求項177記載の方法。

【請求項179】 前記前処理行程は画素画像データの画質を向上する行程を含む請求項178記載の方法。

【請求項180】 更に、文脈チェックを含み、前記認識行程で認識された文字を後処理する行程を備える請求項176記載の方法。

【請求項181】 前記第1の切出行程はラインセグメントをとびとびに処理する行程を含む請求項176記載の方法。

【請求項182】 前記追加の切出行程は、画素画像データの文字について特徴が既知か否かに対応して選択可能である請求項176記載の方法。

【請求項183】 前記追加の切出行程では既知の文字セットの統計に基づいて切り出しを行う請求項182記載の方法。

40 【請求項184】 前記追加の切出行程では画素密度の回転投影に対応して切り出しを行う請求項182記載の方法。

【請求項185】 画素密度の回転投影は画素密度の垂直投影に基づいて決められた角度で得られる請求項184記載の方法。

【請求項186】 更に、前記第1の切出行程と前記少なくとも1つの追加の切出行程の間に中間の切出行程を備える請求項176記載の方法。

50 【請求項187】 前記中間の切出行程はラインセグメ

ント内の重なった画素画像の輪郭を追跡する行程を含む請求項186記載の方法。

【請求項188】 更に、文字が間違っ切り出された場合に、前記中間の切出行程で切り出された文字を再接続する行程を備える請求項187記載の方法。

【請求項189】 前記認識行程は、前記追加の切出行程の前に前記第1の切出行程で切り出された文字を認識し、前記追加の切出行程は前記認識行程で認識されない文字に対して行なわれる請求項188記載の方法。

【請求項190】 更に、文字が前記認識行程で認識されない場合に、前記少なくとも1つの追加の切出行程で切り出された文字を再接続する行程を備える請求項189記載の方法。

【請求項191】 更に、非テキスト要素の内部で白輪郭を追跡する行程を備える請求項176記載の方法。

【請求項192】 更に、非テキスト要素に識別子を付加する行程を備える請求項191記載の方法。

【請求項193】 非テキスト要素内に含まれる白輪郭の数に基づいて非テキスト要素を表と見なす請求項192記載の方法。

【請求項194】 更に、非格子状配置の白輪郭を再接続する行程を備え、非テキスト要素には前記再接続行程での再接続率に基づいて表識別子が付加される請求項192記載の方法。

【請求項195】 更に、非テキスト要素内の白輪郭の充填率を算出する行程を備え、充填率が高くない場合に非テキスト要素を表識別子を付加する請求項192記載の方法。

【請求項196】 画素画像データでラインの位置を割り出す装置であって、画像データの横方向に伸びた部分の全域で画像密度の水平投影を得る手段と、水平投影に基づいて、画像密度が投影された部分の幅を減少するか否かを決定する決定行程と、水平投影に基づいて、ラインの位置を割り出す割出行程とを備える装置。

【請求項197】 前記決定手段は、実質的に空白の部分で水平投影を調べる手段と、実質的に空白の部分が見付からない場合に幅を減少する減少手段とを備える請求項196記載の装置。

【請求項198】 前記調べる手段は間隔の小さい非空白の部分に接続する接続手段を含む請求項197記載の装置。

【請求項199】 前記水平投影を得る手段と決定手段とは、順に繰り返して使用される請求項196記載の装置。

【請求項200】 前記割出手段では、水平投影の実質的に空白の領域間の画像データのライン位置を、ラインとして割り出す請求項196記載の装置。

【請求項201】 更に、画素画像データを前処理する

前処理手段を備える請求項196記載の装置。

【請求項202】 段に分離可能な画素画像データ内でラインの位置を割り出す装置であって、画像データの少なくとも1つの横方向に伸びる段の全域で、画像密度の水平投影を得る手段と、前記手段で得られた水平投影に基づいて、画像密度が投影される段の数を増加するか否かを決定する決定手段と、

前記水平投影に基づいてデジタル画像データ内でラインの位置を割り出す割出手段とを備える装置。

【請求項203】 前記決定手段は、実質的に空白の部分で水平投影を調べる手段と、実質的に空白の部分が見付からない場合に段の数を増加する増加手段とを備える請求項202記載の装置。

【請求項204】 前記調べる手段は間隔の小さい非空白の部分に接続する連結手段を含む請求項203記載の装置。

【請求項205】 前記水平投影を得る手段と決定手段とは、順に繰り返して使用される請求項202記載の装置。

【請求項206】 前記割出手段では、水平投影の実質的に空白の領域間の画像データのライン位置を、ラインとして割り出す請求項202記載の装置。

【請求項207】 更に、画像データの隣接する段のラインが重なるか否かを定める第2の決定手段を備える請求項202記載の装置。

【請求項208】 前記第2の決定手段は、画像データの隣接する段のブロックが重なるか否かを決定するスケール不変の決定手段から成る請求項207記載の装置。

【請求項209】 更に、画素画像データを前処理する前処理手段を備え、前記前処理手段は画像縮小手段を含む請求項202記載の装置。

【請求項210】 前記画像縮小は水平方向と垂直方向とで異なる請求項209記載の装置。

【請求項211】 画素画像データのラインセグメントから文字画像を切り出す装置であって、ラインセグメントから非接触で重なっていない文字を切り出す第1の切出手段と、ラインセグメントから接触文字を切り出す第2の切出手段とを備える装置。

【請求項212】 更に、高解像度の画像セグメントを低解像度に圧縮するように画像データのイメージセグメントを前処理する前処理手段を備える請求項211記載の装置。

【請求項213】 垂直方向と水平方向とで異なる圧縮率が使われる請求項212記載の装置。

【請求項214】 更に、前記第1の切出手段で切り出された文字を認識処理する認識手段を備え、前記第2の切出手段は前記認識手段で認識されなかった文字に対してのみ使用される請求項212記載の装置。

【請求項215】更に、非接触で重なった文字データをラインセグメントから中間で切り出す第3の切出手段を備える請求項214記載の装置。

【請求項216】前記第3の切出手段は、非接触で重なった文字を画素データの輪郭を追跡する輪郭追跡手段を含む請求項215記載の装置。

【請求項217】前記第3の切出手段は、間違って切り出された文字を再接続する再接続手段を含む請求項215記載の装置。

【請求項218】更に、前記認識手段で認識されなかった文字を再接続する第2の再接続手段を備える請求項214記載の装置。

【請求項219】前記第1の切出手段は、前記ラインセグメントを非空白の画素データが見付かるまでとびとびに処理する手段を含む請求項211記載の装置。

【請求項220】更に、前記とびとびの処理手段で非空白の画素が見付かった場合に、緻密に前後をサーチする検索手段を備える請求項219記載の装置。

【請求項221】前記第2の切出手段は非垂直切り出しによって接触文字を切り出す請求項213記載の装置。

【請求項222】非垂直切り出しの角度は画素密度の垂直投影特性に対応して決められる請求項221記載の装置。

【請求項223】更に、垂直投影特性に基づいて少なくとも1つの回転された投影特性を得る手段を備えることを特徴とする請求項222記載の装置。

【請求項224】更に、前記垂直投影特性の角度に近接する角度で複数の回転された投影特性を得る手段を備える請求項223記載の装置。

【請求項225】非垂直切り出しの角度は、前記前記回転投影特性を得る手段と垂直投影特性を得る手段とから得られた投影特性の最小値に基づいて得られる請求項224記載の装置。

【請求項226】前記第2の切出手段は、前記ラインセグメント内の文字の特徴が既知かどうかに基づいて選択可能である請求項211記載の装置。

【請求項227】前記第2の切出手段は、前記文字セット内の文字データのスペースの統計に基づいて切り出す請求項226記載の装置。

【請求項228】更に、小さな文字のセットを再接続する第3の再接続手段を備える請求項227記載の装置。

【請求項229】文字画素データのラインセグメント内で接触文字を切り出す装置であって、ラインセグメント内で画素データの垂直投影特性を得る手段と、

垂直投影特性の最小値から隣接する最大値へ方向から少なくとも1つの角度を算出する第1の算出手段と、前記第1の算出手段で算出された角度に基づいて、少な

くとも1つの回転された投影特性を算出する第2の算出手段と、

前記回転投影特性の最小値に対応する位置で前記回転投影特性の角度にラインセグメントを切る切断手段とを備える装置。

【請求項230】前記第1の算出手段は、前記垂直投影特性を複数のしきい値と比較して前記最小値と隣接する最大値との位置を決定する比較手段を含む請求項229記載の装置。

【請求項231】更に、最小値と隣接する最大値とが見付からない場合に、前記しきい値を増加させる増加手段を備える請求項230記載の装置。

【請求項232】最小値は、少なくとも一方の側が第2しきい値より大きい最大値によって囲まれ、第1しきい値より小さな垂直投影特性上の位置として割り出される請求項231記載の装置。

【請求項233】更に、各々が前記第2の算出手段で算出された回転投影特性に近接する角度で算出される複数の回転投影特性を算出する第3の算出手段を備える請求項229記載の装置。

【請求項234】更に、複数の回転投影特性と垂直投影特性との全体の最小値を割り出す割出手段を備え、前記切断手段は最小値の位置に対応する位置で最小値が得られた角度に対応する角度に切る請求項233記載の装置。

【請求項235】画素画像データ内で文字を認識する装置であって、画素画像データから画素ブロックを選別する選別手段と、

少なくとも1つの段全体の画素密度の水平投影に基づいてテキストブロックを少なくとも1つの段に適応的に分配することにより、テキストブロックを画素データのラインに分割する分割手段と、

非接触で重なっていない文字がラインセグメントから切り出される第1の切出手段と、接触文字を切り出す少なくとも1つの追加の切出手段とを含み、前記分割手段で分割されたラインセグメントから文字を切り出す切出手段と、

前記切出手段で切り出された文字を認識する認識手段と、前記認識手段で認識された文字を前記選別手段で確定された順に対応して格納する格納手段とを備える装置。

【請求項236】更に、画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する変換手段を含む画素画像データの入力手段を備える請求項235記載の装置。

【請求項237】更に、画素画像データが前処理される前処理手段を備える請求項235記載の装置。

【請求項238】前記前処理手段は画像圧縮手段を含む請求項237記載の装置。

【請求項239】 前記前処理手段は画素画像データの画質を向上するフィルタ手段を含む請求項237記載の装置。

【請求項240】 更に、前記認識手段で認識された文字を後処理する後処理手段を備える請求項235記載の装置。

【請求項241】 前記後処理手段は文脈チェック手段を含む請求項240記載の装置。

【請求項242】 前記第1の切出手段はラインセグメントをとびとびに処理する手段を含む請求項235記載の装置。

【請求項243】 前記追加の切出手段は、画素画像データの文字について特徴が既知か否かに対応して選択可能である請求項235記載の装置。

【請求項244】 前記追加の切出手段は既知の文字セットの統計に基づいて切り出しを行う請求項243記載の装置。

【請求項245】 前記追加の切出手段は画素密度の回転投影に対応して切り出しを行う請求項243記載の装置。

【請求項246】 画素密度の回転投影は画素密度の垂直投影に基づいて決められた角度で得られる請求項245記載の装置。

【請求項247】 更に、前記第1の切出手段の使用後で前記少なくとも1つの追加の切出手段の使用前に、中間の階層を切り出す中間の切出手段を備える請求項245記載の装置。

【請求項248】 前記中間の切出手段はラインセグメント内の重なった画素画像の輪郭を追跡する輪郭追跡手段を含む請求項247記載の装置。

【請求項249】 更に、文字が間違っって切り出された場合に、前記中間の切出手段で切り出された文字を再接続する再接続手段を備える請求項248記載の装置。

【請求項250】 前記認識手段は、前記追加の切出手段の前に前記第1の切出手段で切り出された文字を認識し、前記追加の切出手段は前記認識手段で認識されない文字に対して行なわれる請求項249記載の装置。

【請求項251】 更に、文字が前記認識手段で認識されない場合に、前記少なくとも1つの追加の切出手段で切り出された文字を再接続する第2の再接続手段を備える請求項250記載の装置。

【請求項252】 画素画像データから文字のテキストファイルを形成する装置であって、画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する変換手段を含む画素画像データの入力手段と、画素画像データのブロックがテキストブロックか非テキストブロックかに対応して画素画像データのブロックを選別する選別手段と、テキスト及び非テキストブロックの階層的木構造を得る

手段と、

少なくとも1つの段の全域の画素密度の水平投影に基づいて、テキストブロックを前記少なくとも1つの段に適応的に分配することにより、テキストブロックを画素画像データのラインに分割する分割手段と、

非接触で重なっていない文字をラインセグメントから切り出す第1の切出手段と、接触文字を切り出す少なくとも1つの追加の切出手段とを含み、前記分割手段で分割されたラインから文字を切り出す切出手段と、

前記切出手段で切り出された文字を認識し、前記認識に基づいて文字コードを割り付ける認識手段と、前記階層的木構造により確立された順に対応して文字コードが格納されるコンピュータのテキストファイルとを備える装置。

【請求項253】 更に、画素画像データが前処理される前処理手段を備える請求項252記載の装置。

【請求項254】 前記前処理手段は画像圧縮手段を含む請求項253記載の装置。

【請求項255】 前記前処理手段は画素画像データの画質を向上する画質向上手段を含む請求項253記載の装置。

【請求項256】 更に、文脈チェック手段を含み、前記認識手段で認識された文字を後処理する後処理手段を備える請求項252記載の装置。

【請求項257】 前記第1の切出手段はラインセグメントをとびとびに処理する手段を含む請求項252記載の装置。

【請求項258】 前記追加の切出手段は、画素画像データの文字について特徴が既知か否かに対応して選択可能である請求項252記載の装置。

【請求項259】 前記追加の切出手段は既知の文字セットの統計に基づいて切り出しを行う請求項258記載の装置。

【請求項260】 前記追加の切出手段は画素密度の回転投影に対応して切り出しを行う請求項258記載の装置。

【請求項261】 画素密度の回転投影は画素密度の垂直投影に基づいて決められた角度で得られる請求項260記載の装置。

【請求項262】 更に、前記第1の切出手段と前記少なくとも1つの追加の切出手段の間に中間の階層を切り出す中間の切出手段を備える請求項252記載の装置。

【請求項263】 前記中間の切出手段はラインセグメント内の重なった画素画像の輪郭を追跡する輪郭追跡手段を含む請求項262記載の装置。

【請求項264】 更に、文字が間違っって切り出された場合に、前記中間の切出手段で切り出された文字を再接続する再接続手段を備える請求項263記載の装置。

【請求項265】 前記認識手段は、前記追加の切出手段の前に前記第1の切出手段で切り出された文字を認識

し、前記追加の切出手段は前記認識手段で認識されない文字に対して行なわれる請求項264記載の装置。

【請求項266】 更に、文字が前記認識手段で認識されない場合に、前記少なくとも1つの追加の切出手段で切り出された文字を再接続する第2の再接続手段を備える請求項265記載の装置。

【請求項267】 画素画像データから文字のテキストファイルを形成する装置であって、

画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する変換手段を含み、画素画像データを入力する入力手段と、画素画像データ内で連結要素の輪郭を追跡する輪郭追跡手段と、追跡された連結要素のサイズに基づいて追跡された連結要素がテキスト要素を含むか非テキスト要素を含むから決める決定手段と、隣接するテキスト要素の近接関係に基づいてテキスト要素を選択的に横方向に連結してテキストラインを形成する第1の連結手段と、隣接するテキストラインの近接関係とテキストライン間の非テキスト要素の位置とに基づいてテキストラインを選択的に縦方向に連結する第2の連結手段とを含み、画素画像データのブロックを選別する選別手段と、前記輪郭追跡手段で追跡された連結要素に基づいて、階層の木構造を形成する形成手段と、少なくとも1つの段の全域での画素密度の水平投影に基づいて、テキストブロックを前記少なくとも1つの段に適応的に分配することによって、テキストブロックを画素データのラインに分割する分割手段と、非接触で重ならない文字をラインセグメントから切り出す第1の切出手段と、接触文字を切り出す少なくとも1つの追加の切出手段とを含み、前記分割手段で分割されたラインから文字を切り出す切出手段と、前記文字切出手段で切り出された文字を認識し、前記認識に基づいて文字コードを割り付ける認識手段と、階層の木構造により確定された順に対応して文字コードをコンピュータのテキストファイルに格納する格納手段とを備える装置。

【請求項268】 更に、画素画像データが前処理される前処理手段を備える請求項267記載の装置。

【請求項269】 前記前処理手段は画像圧縮手段を含む請求項268記載の装置。

【請求項270】 前記前処理手段は画素画像データの画質を向上する画質向上手段を含む請求項269記載の装置。

【請求項271】 更に、文脈チェック手段を含み、前記認識手段で認識された文字を後処理する後処理手段を備える請求項267記載の装置。

【請求項272】 前記第1の切出手段はラインセグメントをとびとびに処理する手段を含む請求項267記載の装置。

【請求項273】 前記追加の切出手段は画素画像デー

タの文字について特徴が既知か否かに対応して選択可能である請求項267記載の装置。

【請求項274】 前記追加の切出手段は既知の文字セットの統計に基づいて切り出しを行う請求項273記載の装置。

【請求項275】 前記追加の切出手段は画素密度の回転投影に対応して切り出しを行う請求項273記載の装置。

【請求項276】 画素密度の回転投影は画素密度の垂直投影に基づいて決められた角度で得られる請求項275記載の装置。

【請求項277】 更に、前記第1の切出手段の後で前記少なくとも1つの追加の切出手段の前に中間の階層を切り出す中間の切出手段を備える請求項267記載の装置。

【請求項278】 前記中間の切出手段はラインセグメント内の重なった画素画像の輪郭を追跡する輪郭追跡手段を含む請求項277記載の装置。

【請求項279】 更に、文字が間違って切り出された場合に、前記中間の切出手段で切り出された文字を再接続する第1の再接続手段を備える請求項278記載の装置。

【請求項280】 前記認識手段は、前記追加の切出手段の前に前記第1の切出手段で切り出された文字を認識し、前記追加の切出手段は前記認識手段で認識されない文字に対して行なわれる請求項279記載の装置。

【請求項281】 更に、文字が前記認識手段で認識されない場合に、前記少なくとも1つの追加の切出手段で切り出された文字を再接続する第2の再接続手段を備える請求項280記載の装置。

【請求項282】 更に、非テキスト要素の内部で白輪郭を追跡する手段を備える請求項267記載の装置。

【請求項283】 更に、非テキスト要素に識別子を付加する付加手段を備える請求項282記載の装置。

【請求項284】 非テキスト要素内に含まれる白輪郭の数に基づいて非テキスト要素を表と見なす請求項283記載の装置。

【請求項285】 更に、非格子状配置の白輪郭を再接続する手段を備え、非テキスト要素には前記再接続手段での再接続率に基づいて表の識別子が付加される請求項283記載の装置。

【請求項286】 更に、非テキスト要素内の白輪郭の充填率を算出する算出手段を備え、充填率が低い場合に非テキスト要素に表の識別子を付加する請求項283記載の装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、文字認識方法及び装置に関わるものであり、特に認識処理に先立って、画像データに基づいて画像データブロックを分類・選別する方

し、前記追加の切出手段は前記認識手段で認識されない文字に対して行なわれる請求項264記載の装置。

【請求項266】 更に、文字が前記認識手段で認識されない場合に、前記少なくとも1つの追加の切出手段で切り出された文字を再接続する第2の再接続手段を備える請求項265記載の装置。

【請求項267】 画素画像データから文字のテキストファイルを形成する装置であって、

画素画像データが2値の画素画像データでない場合に、画素画像データを2値の画素画像データに変換する変換手段を含み、画素画像データを入力する入力手段と、画素画像データ内で連結要素の輪郭を追跡する輪郭追跡手段と、追跡された連結要素のサイズに基づいて追跡された連結要素がテキスト要素を含むか非テキスト要素を含むから決める決定手段と、隣接するテキスト要素の近接関係に基づいてテキスト要素を選択的に横方向に連結してテキストラインを形成する第1の連結手段と、隣接するテキストラインの近接関係とテキストライン間の非テキスト要素の位置とに基づいてテキストラインを選択的に縦方向に連結する第2の連結手段とを含み、画素画像データのブロックを選別する選別手段と、前記輪郭追跡手段で追跡された連結要素に基づいて、階層的木構造を形成する形成手段と、少なくとも1つの段の全域での画素密度の水平投影に基づいて、テキストブロックを前記少なくとも1つの段に適応的に分配することによって、テキストブロックを画素データのラインに分割する分割手段と、非接触で重ならない文字をラインセグメントから切り出す第1の切出手段と、接触文字を切り出す少なくとも1つの追加の切出手段とを含み、前記分割手段で分割されたラインから文字を切り出す切出手段と、前記文字切出手段で切り出された文字を認識し、前記認識に基づいて文字コードを割り付ける認識手段と、階層的木構造により確定された順に対応して文字コードをコンピュータのテキストファイルに格納する格納手段とを備える装置。

【請求項268】 更に、画素画像データが前処理される前処理手段を備える請求項267記載の装置。

【請求項269】 前記前処理手段は画像圧縮手段を含む請求項268記載の装置。

【請求項270】 前記前処理手段は画素画像データの画質を向上する画質向上手段を含む請求項269記載の装置。

【請求項271】 更に、文脈チェック手段を含み、前記認識手段で認識された文字を後処理する後処理手段を備える請求項267記載の装置。

【請求項272】 前記第1の切出手段はラインセグメントをとびとびに処理する手段を含む請求項267記載の装置。

【請求項273】 前記追加の切出手段は画素画像デー

タの文字について特徴が既知か否かに対応して選択可能である請求項267記載の装置。

【請求項274】 前記追加の切出手段は既知の文字セットの統計に基づいて切り出しを行う請求項273記載の装置。

【請求項275】 前記追加の切出手段は画素密度の回転投影に対応して切り出しを行う請求項273記載の装置。

【請求項276】 画素密度の回転投影は画素密度の垂直投影に基づいて決められた角度で得られる請求項275記載の装置。

【請求項277】 更に、前記第1の切出手段の後で前記少なくとも1つの追加の切出手段の前に中間の階層を切り出す中間の切出手段を備える請求項267記載の装置。

【請求項278】 前記中間の切出手段はラインセグメント内の重なった画素画像の輪郭を追跡する輪郭追跡手段を含む請求項277記載の装置。

【請求項279】 更に、文字が間違って切り出された場合に、前記中間の切出手段で切り出された文字を再接続する第1の再接続手段を備える請求項278記載の装置。

【請求項280】 前記認識手段は、前記追加の切出手段の前に前記第1の切出手段で切り出された文字を認識し、前記追加の切出手段は前記認識手段で認識されない文字に対して行なわれる請求項279記載の装置。

【請求項281】 更に、文字が前記認識手段で認識されない場合に、前記少なくとも1つの追加の切出手段で切り出された文字を再接続する第2の再接続手段を備える請求項280記載の装置。

【請求項282】 更に、非テキスト要素の内部で白輪郭を追跡する手段を備える請求項267記載の装置。

【請求項283】 更に、非テキスト要素に識別子を付加する付加手段を備える請求項282記載の装置。

【請求項284】 非テキスト要素内に含まれる白輪郭の数に基づいて非テキスト要素を表と見なす請求項283記載の装置。

【請求項285】 更に、非格子状配置の白輪郭を再接続する手段を備え、非テキスト要素には前記再接続手段での再接続率に基づいて表の識別子が付加される請求項283記載の装置。

【請求項286】 更に、非テキスト要素内の白輪郭の充填率を算出する算出手段を備え、充填率が高くない場合に非テキスト要素に表の識別子を付加する請求項283記載の装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、文字認識方法及び装置に関わるものであり、特に認識処理に先立って、画像データに基づいて画像データブロックを分類・選別する方

法や装置に関するものである。例えば、画像データがテキスト画像データであるか、中間調（あるいはグレースケール）画像、線画及びフレームなどのような非テキスト画像であるかに応じて、画像データブロックを選別して分類するものである。また、本発明は、認識処理への準備として、テキストブロックからテキストラインを識別して分割し、次にテキストライン中の個々の文字を識別してライン中の他の文字から切り出す方法や装置にも関するものである。

【0002】尚、本出願は付録のコンピュータプログラム（図35～図247参照）とともに提出される。本明細書の開示の一部には、著作権保護に関係するものが含まれている。著作権者は、特許・商標庁の特許ファイルあるいは記録なので本出願の開示文書がファクシミリ再生されることに対しては異議を唱えないが、そうでない場合にはすべての著作権を留保するものとする。

【0003】

【従来の技術】近年、テキストデータ画像を分析してテキストデータ中の個々の文字を認識し、認識された文字に対応してコンピュータが理解可能な文字コードファイルを生成することが可能になってきた。ここで得られたファイルは、ワードプロセッシングやデータプロセッシングのプログラムで操作を加えることができる。以下「文字認識システム」と呼ぶことにするこのようなシステムの利点は、テキストデータを再タイプしたり、あるいは再入力する必要がないことである。例えば、ファクシミリで送られてきた文書や、マイクロフィルムや複写機で再生された文書を文字認識して、文書中の文字や数字を文字コード（例えばアスキーコードコード）からなるコンピュータテキストファイルとして生成することにより、文書の再タイプや再入力をせずに、文書をワードプロセッシングやデータプロセッシングすることができる。

【0004】文字認識すべき文書にはさまざまな種類の画像が含まれることが多く、すべてが認識可能というわけではない。例えば、テキスト画像データの認識は現在でも可能であるが、非テキスト画像データの認識は不可能である。一般に、文字認識すべき文書には、テキスト画像データブロックと中間調画像、線画及び線などの非テキスト画像データブロックとが含まれる。また、文書には、フレームで囲まれていたり囲まれていなかったりする表や表形式のデータなども含まれる。したがって、文字認識処理に先立って、ブロック中の画像データの種類の応じて文書中の個々のブロックを分類し、画像データからテキストタイプのブロックを選別することが必要となる。

【0005】図32は典型的な文書の1ページの例を示す図である。図32において、文書ページ401は2段組のフォーマットで構成されている。ページには、タイトルに適した大きなフォントサイズのテキスト情報が含

まれるタイトルブロック402、テキストデータラインを含むテキストブロック404、テキストではないグラフィックス画像を含むグラフィックスブロック405、テキストや数字情報の表を含む表ブロック406、小さなサイズのテキストデータでグラフィックスブロックあるいは表ブロックに対する見だしが含まれる見だしブロック407が存在する。各情報ブロックはブロック内の情報の種別に応じて分類され、この分類結果に基づいてブロックが分割される。

【0006】画像データからテキスト型のブロックを検出するため、従来は、画像データ中の黒画素を隣接する1つあるいは複数の白画素にまで水平方向・垂直方向に膨張させて、水平方向・垂直方向に画素画像データを塗りつぶす手法を用いていた。このような塗りつぶし法の欠点としては、適切な塗りつぶしパラメータを選択するために、前もってテキスト型の画像データの特徴（例えばフォントサイズなど）を知らなければならないことが挙げられる。また、塗りつぶしパラメータの僅かな変更が、選別結果の大幅な変化につながるという問題もある。さらに、塗りつぶし法では、必ずしも原文書の内部構造が保存されるとは限らない。例えば、塗りつぶし法では、2段組の原文書が1段に塗りつぶされてしまうことがある。このような場合には、テキストデータの格納順序がばらばらになってしまい、原テキストの正確な再生が不可能という問題が生じる。この他、テキスト型のデータが非テキスト型のデータまで塗りつぶして、全領域をテキスト型のデータと誤って解釈してしまう場合も、塗りつぶし法ではしばしば生じる。

【0007】ブロック選別に続いて、文書中の文字ごとに文字認識処理を行い、文字に対応するコンピュータコードを得る。ここで、文字ブロックから個々の文字を得るには、2つの処理を行う。第1段の処理では、タイトルブロック202、テキストブロック204、見だしブロック207などの各テキストブロック中の個々のラインが、テキストブロック中の他のラインから分割される。一般に、ライン分割は、各ブロック中の画素密度を水平投影してライン間のギャップを識別することで行われる。すなわち、図33Aに示すように、テキストブロック404は間にギャップ412を有するテキストラインからなる。画素密度の水平投影414は、ブロック404の各ライン上の黒画素数の総和として求められる。テキストラインは投影414中で密度が非ゼロの領域に対応し、テキストライン間のギャップは投影414中の密度がゼロの領域に対応する。したがって、投影密度に基づいて、テキストラインを各ラインごとに分割することができる。

【0008】第2段の処理では、分割されたテキストライン中の個々の文字がテキストライン中の他の文字から切り出される。すなわち、図34Aに示すように、テキストライン411には個々の文字が含まれる。テキスト

法や装置に関するものである。例えば、画像データがテキスト画像データであるか、中間調（あるいはグレースケール）画像、線画及びフレームなどのような非テキスト画像であるかに応じて、画像データブロックを選別して分類するものである。また、本発明は、認識処理への準備として、テキストブロックからテキストラインを識別して分割し、次にテキストライン中の個々の文字を識別してライン中の他の文字から切り出す方法や装置にも関するものである。

【0002】尚、本出願は付録のコンピュータプログラム（図35～図247参照）とともに提出される。本明細書の開示の一部には、著作権保護に関係するものが含まれている。著作権者は、特許・商標庁の特許ファイルあるいは記録なので本出願の開示文書がファクシミリ再生されることに対しては異議を唱えないが、そうでない場合にはすべての著作権を留保するものとする。

【0003】

【従来の技術】近年、テキストデータ画像を分析してテキストデータ中の個々の文字を認識し、認識された文字に対応してコンピュータが理解可能な文字コードファイル（20）を生成することが可能になってきた。ここで得られたファイルは、ワードプロセッシングやデータプロセッシングのプログラムで操作を加えることができる。以下「文字認識システム」と呼ぶことにするこのようなシステムの利点は、テキストデータを再タイプしたり、あるいは再入力する必要がないことである。例えば、ファクシミリで送られてきた文書や、マイクロフィルムや複写機で再生された文書を文字認識して、文書中の文字や数字を文字コード（例えばアスキーコードコード）からなるコンピュータテキストファイルとして生成することにより、文書の再タイプや再入力をせずに、文書をワードプロセッシングやデータプロセッシングすることができる。

【0004】文字認識すべき文書にはさまざまな種類の画像が含まれることが多く、すべてが認識可能というわけではない。例えば、テキスト画像データの認識は現在でも可能であるが、非テキスト画像データの認識は不可能である。一般に、文字認識すべき文書には、テキスト画像データブロックと中間調画像、線画及び線などの非テキスト画像データブロックとが含まれる。また、文書（40）には、フレームで囲まれていたり囲まれていなかったりする表や表形式のデータなども含まれる。したがって、文字認識処理に先立って、ブロック中の画像データの種類に応じて文書中の個々のブロックを分類し、画像データからテキストタイプのブロックを選別することが必要となる。

【0005】図32は典型的な文書の1ページの例を示す図である。図32において、文書ページ401は2段組のフォーマットで構成されている。ページには、タイトルに適した大きなフォントサイズのテキスト情報が含

まれるタイトルブロック402、テキストデータラインを含むテキストブロック404、テキストではないグラフィックス画像を含むグラフィックスブロック405、テキストや数字情報の表を含む表ブロック406、小さなサイズのテキストデータでグラフィックスブロックあるいは表ブロックに対する見だしが含まれる見だしブロック407が存在する。各情報ブロックはブロック内の情報の種別に応じて分類され、この分類結果に基づいてブロックが分割される。

【0006】画像データからテキスト型のブロックを検出するため、従来は、画像データ中の黒画素を隣接する1つあるいは複数の白画素にまで水平方向・垂直方向に膨張させて、水平方向・垂直方向に画素画像データを塗りつぶす手法を用いていた。このような塗りつぶし法の欠点としては、適切な塗りつぶしパラメータを選択するために、前もってテキスト型の画像データ中の特徴（例えばフォントサイズなど）を知らなければならないことが挙げられる。また、塗りつぶしパラメータの僅かな変更が、選別結果の大幅な変化につながるという問題もある。さらに、塗りつぶし法では、必ずしも原文書の内部構造が保存されるとは限らない。例えば、塗りつぶし法では、2段組の原文書が1段に塗りつぶされてしまうことがある。このような場合には、テキストデータの格納順序がばらばらになってしまい、原テキストの正確な再生が不可能という問題が生じる。この他、テキスト型のデータが非テキスト型のデータまで塗りつぶして、全領域をテキスト型のデータと誤って解釈してしまう場合も、塗りつぶし法ではしばしば生じる。

【0007】ブロック選別に続いて、文書中の文字ごとに文字認識処理を行い、文字に対応するコンピュータコードを得る。ここで、文字ブロックから個々の文字を得るには、2つの処理を行う。第1段の処理では、タイトルブロック202、テキストブロック204、見だしブロック207などの各テキストブロック中の個々のラインが、テキストブロック中の他のラインから分割される。一般に、ライン分割は、各ブロック中の画素密度を水平投影してライン間のギャップを識別することで行われる。すなわち、図33Aに示すように、テキストブロック404は間にギャップ412を有するテキストラインからなる。画素密度の水平投影414は、ブロック404の各ライン上の黒画素数の総和として求められる。テキストラインは投影414中で密度が非ゼロの領域に対応し、テキストライン間のギャップは投影414中の密度がゼロの領域に対応する。したがって、投影密度に基づいて、テキストラインを各ラインごとに分割することができる。

【0008】第2段の処理では、分割されたテキストライン中の個々の文字がテキストライン中の他の文字から切り出される。すなわち、図34Aに示すように、テキストライン411には個々の文字が含まれる。テキスト

ラインにおいて各文字を他の文字から切り出すために、ラインセグメント411の各列ごとに黒画素数の垂直方向の総和を求め、画素密度の垂直投影416を得る。文字415は投影416中で密度が非ゼロの領域に対応し、文字間のギャップは投影416中で密度がゼロの領域に対応する。こうして、個々の文字がラインセグメント中の他の文字から切り出される。

【0009】このような処理においてはいくつかの問題点が存在する。例えば、文書が斜めに画像スキャナに入力され、図33Bに示すように傾斜角 θ sで画像メモリに格納される場合が多いが、このような場合には、第1ライン418のテキストが第2ライン419のテキストとライン420で重なってしまうため、必ずしもライン分割が可能であるとは限らない。すなわち、画素密度の水平投影421は非ゼロ値のみとなり、ゼロ値が存在しないためライン間のギャップの検出が不可能になる。

【0010】このような問題点に鑑みて、テキストブロック404を図33Cに示すような複数の列422、424に分割して、各列ごとに独立に水平投影を得る手法が用いられている。すなわち、図33Cに示すように、水平投影422aは列422に対応し、水平投影424aは列424に対応するものである。各列でテキストラインが重なっていなければ、各列中のテキストラインの検出が可能となる。図33Cでは2列に分割する例を示しているが、一般には5列から10列に分割し、テストで最大傾斜角 θ sまでの傾きを有していても、ブロック中の他のラインから個々のラインの分割が可能であるようにしている。

【0011】しかし、各列ごとに水平画素投影を得なければならない、またこのようにして得た各水平画素投影も個々に処理しなければならないため、ライン分割処理がきわめて時間のかかる処理となる。また、最大傾斜角 θ sの文書にまで対応しなければならないため、傾斜角が小さくて1列あるいは数列への分割で十分である多くの文書に対しても、すべての列を処理しなければならない、この点でも時間のかかる処理となっている。上記のような2段階処理のもう1つの問題は、個々の文字をラインセグメント中の他の文字から切り出す第2段の処理において生じる。

【0012】すなわち、文字間に垂直な空間が存在する場合には、図34Aに示したような処理で良好な結果が得られるが、文字が垂直方向に重なっていた場合や2つあるいは複数の文字が接触している場合には、文字分割処理を適切に行うことができない。イタリックフォントの場合や、複数回の複写処理やファクシミリ送信によって画像品質が劣化している場合など、このような状況に陥る可能性は高い。例えば、図34Bに示すように、イタリックテキストでは単語「Satisfy」中の文字「f」と「y」とが垂直方向に重なっており、画素密度の垂直投影425が文字間でゼロにならないことがあ

る。そのため、文字「f」と「y」とを分離することはできない。また、文字「t」と「i」も接触しているため、同様にこれら2つの文字の分離も不可能である。

【0013】

【発明が解決しようとしている課題】本発明は、前記従来の欠点を除去し、高速かつ正確に文書上の文字を認識してテキストファイルを作成する文字認識方法及び装置を提供する。また、高速かつ正確に文書上のテキストと非テキストとを選別して、テキストブロックを割り出す方法及び装置を提供する。

【0014】また、高速かつ正確に文書上の非テキストを分類する方法及び装置を提供する。また、高速かつ正確にテキストブロックからテキストラインを分割する方法及び装置を提供する。また、高速かつ正確に傾いた文書のテキストブロックからテキストラインを分割する方法及び装置を提供する。

【0015】また、高速かつ正確にテキストラインから文字を切り出す方法及び装置を提供する。また、不適切に切り出された文字の再切り出しが可能な方法及び装置を提供する。

【0016】

【課題を解決するための手段】この課題を解決するために、本発明の文字認識方法及び装置は以下の構成を含む。画素画像データが入力され、2値の画素画像データでない場合は2値の画素画像データに変換される。

【0017】画素画像データのブロックが、画素画像データ内で連結要素の輪郭を追跡し、連結要素がテキスト要素を含むか非テキスト要素を含むかを連結要素のサイズに基づいて決定し、隣接するテキスト要素の近接状態に基づいてテキスト要素を選択的に連結してテキストラインを形成し、隣接するテキストラインの近接状態とテキストライン間の非テキスト要素の位置とに基づいてテキストラインを選択的に連結することにより、選別される。階層的木構造が連結要素に基づいて形成される。

【0018】テキストブロックは、段の全域での画素密度の水平投影に基づいてテキストブロックを少なくとも1つのラインに分割することで、テキストラインに分割され、非接触で重なっていない文字を切り出す第1の切り出しと接触文字間を切り離す第2の切り出しとの2回の切り出しで、文字が分割されたラインから切り出される。切り出された文字は認識され、認識に基づいて文字コードが割り当てられる。文字コードは階層的木構造による確定された順でコンピュータのテキストファイル内に格納される。

【0019】必要ならば、非テキスト要素は階層的木構造により確定された順に対応して格納された文字コードの間に散在してもよい。画素画像データは例えば画像圧縮や画質向上等の前処理をされ、認識された文字は例えば文脈チェック等の後処理をされてもよい。非テキスト要素の特徴に基づいて、非テキスト要素に識別子が付与

される。例えば、非テキスト要素の内部で白輪郭の追跡が行なわれ、非格子状配置の白輪郭が再連結されて、白輪郭の充填率が計算されて、白輪郭の数、非格子状配置の白輪郭の再連結率、あるいは白輪郭の充填率等に基づいて、非テキスト要素には表の識別子が付与される。

【0020】非接触で重なっていない文字の切り出しは、ラインセグメントへのとびとびの処理で達成され、接触文字の切り離しは、接触文字間の空間に関する情報が既知か否かに対応して達成される。空間に関する情報が既知の場合は、空間の統計値に基づいて切り出しが達成される。空間に関する情報が無い場合、接触文字間の切り離しは、回転された投影によって決められた角度と位置で接触文字を斜めに切り離すように、画素密度の回転投影に対応して達成される。不適切に切り出された文字は再接続される。

【0021】

【実施例】本発明は、複写機、ファクシミリ、ビデオカメラあるいはスチルビデオカメラ、レーザビームプリンタなどの画像処理装置あるいは画像再生装置などのような、文字認識処理が望まれるさまざまな装置において実施可能である。このような装置においては、文字画像を有する画像を、文字画像が文字認識されるように処理あるいは再生する。この際、認識した文字画像を標準文字セットあるいはフォントに替えて、原文字画像ではなく標準文字を再送あるいは再生することもできる。また、本発明は、汎用コンピュータや、パーソナルコンピュータ、ワードプロセッシングあるいはデータプロセッシング機器などのオフィス機器や、複数のビジネス機器を単一の統合パッケージにまとめた統合オフィス自動機器などにおいても実施可能である。

【0022】図1は、スキャン、ファクシミリ、情報送受信、情報処理（ワードプロセッシングやデータプロセッシングなどの処理をも含む）などの機能を有する統合オフィス自動機器として、本発明の一実施例を示すブロック図である。図1に示す装置において、画像の入力は、ファクシミリ送信による入力、原文書のスキャンによる入力、モデムを介して離れた地点からの入力などで行われる。本実施例によれば、画像中の文字を認識し、認識文字のコンピュータテキストファイルを生成し、装置内のワードプロセッシング、スプレッドシート処理ならびに他の情報処理機能を利用してテキストファイルの修正を行うことができる。この処理に続いて、修正されたテキストファイル（あるいは無修正のテキストファイル）の再送や、音声合成技術を用いてテキストファイルのスピーカや通常の音声電話への「話しかけ」などによる出力が行われる。

【0023】図1において、プログラム可能なマイクロプロセッサなどの中央処理ユニット（CPU）10はバス11に接続されている。また、画像を画素ごとにスキャンして画像メモリ（例えば、以下で示すRAM20）

に蓄積するスキャナ12、デジタルデータを電話線15aを通してアナログ形式で送受信するモデム14、画像データを電話線15bを通して送受信するファクシミリ16（不図示の電話を含むこともある）なども、バス11に接続される。なお、電話線15aと15bとは同一の線であって、不図示のネットワーク制御ユニットで制御されるものでも良い。さらに、バス11には、CPU10によって実行される1つあるいは複数のコンピュータプログラムを蓄えるリードオンリメモリ（ROM）17、認識処理において入力文字と比較する文字の辞書を保持している文字辞書19、入力画像データ、処理画像データならびに画像の構造などの情報を蓄えるランダムアクセスメモリ（RAM）20、文字認識処理において文字の認識結果を出力する出力装置21（ディスクメモリ、スピーカあるいは音声電話線インタフェースを有する音声合成装置など）、装置によって処理された画像を表示するプリンタ/ディスプレイ22、オペレータが装置を操作するためのキーボード24なども接続される。

【0024】なお、ここでは、バス11に接続された装置が協調して統合オフィス自動機器を構成しているが、これらの装置のうちのいくつかあるいはすべてが単独で用いられることも可能なことは明らかであろう。スキャナ12、モデム14及びファクシミリ16は、画像データを装置に入力するための種々の入力手段である。スキャナ12では、原画像はラインごとと画素ごとにスキャンされ、CPU10の制御のもとで画像データの画素はRAM20中の画像メモリにビットマップ形式で蓄えられる。モデム14では、電話線15aを介してアナログ形式で受信した画像データがデジタル画素形式に変換され、デジタル画素データはRAM20中の画像メモリに蓄えられる。ファクシミリ16では、電話線15bを介して修正ハフマンランレンクス符号などの符号形式で受信した圧縮画像データを、ファクシミリ16において従来手法でデジタル画像の画素データに復号して、CPU10によって画像データの画素はRAM20中の画像メモリにビットマップ形式で蓄えられる。ここで、他の画像入力手段を用いることはもちろん可能で、ディスクメモリなどの大容量の蓄積メディアから画像データを取り出したり、ビデオカメラあるいはスチルビデオカメラから入力することもできる。

【0025】ファクシミリ16や出力装置21は、文字認識された画像データを装置から出力するための種々の手段である。ファクシミリ16では、本実施例に基づき認識処理された文字画像が標準文字セットあるいはフォントに変換され、装置から送信される。これにより、例えば、文字画像を含む画像を受信して、文字画像を文字認識し、再送に先立って認識文字を標準文字フォントに変換することで、劣化画像の品質の改善を図ることができる。

【0026】モデム14や出力装置21は、画像データ

で認識された文字をASCIIコードなどで出力あるいは蓄積するための種々の手段である。すなわち、文字は装置（ディスクメモリなど）に蓄えられたり、モデム14を介して離れた地点に転送するために出力される。この際、ASCIIコードをファクシミリ互換のフォーマットに変換するなどの文字変換手段を設ければ、ファクシミリ16を用いることなくモデム14を介して離れた地点のファクシミリ機に転送を行うことができる。

【0027】プリンタ/ディスプレイ22は、本実施例に基づく文字認識処理の流れの監視、ならびに文字認識の各ステップの永久記録の出力および表示を行うための手段である。キーボード24は、図1の装置の動作に対するオペレータの制御を可能にしている。図2は、本実施例の文字認識処理のフローチャートである。図2に示した処理ステップは、プログラムROM17に蓄えられているコンピュータプログラムに基づいてCPU10で実行される。

【0028】ステップS201において、画素単位の画像データ（以下画素画像データ）は装置に入力され、RAM17に蓄えられる。画像データは画像を画素ごとに表現したものである。この際、画素データは2値画素データ、すなわち黒ならびに白画像データであることが好ましい。なお、画像データは、複数のグレイ階調値の1つで各画素が表現される中間調画像であっても、画素色を表す複数ビットワードで各画素が表現されるカラー画像データであっても良い。このような場合、すなわち画素データが2値画素データでない場合には、非2値画素データを2値画素データに変換するためのしきい値処理をRAM20に蓄積する前に行う必要がある。

【0029】ステップS201で入力された画素画像データは、左上隅から右下隅に読み進められるようなポートレート画像であることが好ましい。画像がポートレート画像でない場合、例えばランドスケープ画像である場合には、画素画像データをポートレート型の画像に変換する必要がある。この変換処理は、オペレータがキーボード24から画像データの変換を指示することで行うことができる。

【0030】ステップS202では、画像データの前処理が行われる。一般に、前処理フィルターを用いて画像データの画質を向上し、劣化した文字や画像を改善する。画質向上手法として適切な手法は、1991年10月4日出願の米国出願番号07/771,220に示されている。尚、ステップS202では、精度の劣化は避けられないが、画素画像データ中の複数の画素を削除したり圧縮したりすることで、認識処理の高速化を図ることも可能である。例えば、 $m \times n$ 画素ブロックの画素の平均値を求め、この平均値で $m \times n$ 画素ブロックを代表させることも可能である。

【0031】ステップS203では、ブロック選択を行って各セグメント画像中の画像データの型を調べ、プロ

ック中の情報がテキスト情報か、グラフィックス情報か、線画情報か、画像情報かなどを識別する。また、ブロック選択ステップS203では、画像中の各部分を以下のステップS212で示すような適切な順序で再生することができるように、画像の階層的木構造をも生成する。この階層的木構造には、例えば、2段組の画像データにおいて第1段のテキストから第2段のテキストに飛んで読み進めることのないように、第1段のテキストの後に第2段のテキストを再生できるような情報が含まれる。なお、ステップS203におけるブロック選択処理については、以下でより詳細に説明する。

【0032】RAM20に蓄えられた画像から第1の情報ブロックが選択され、ブロック選択ステップS203で判別されたブロック識別子に応じて、ステップS204において選択されたブロックがテキストブロックであるかどうか調べられる。第1ブロックがテキストブロックでない場合には、処理をステップS205に進め、次のブロックを選択してステップS204に戻る。

【0033】ステップS204においてブロックがテキストブロックである場合には、処理をステップS206に進め、テキストブロックに対してライン分割処理を行う。ライン分割処理では、テキストブロック中の個々のテキストラインをテキストブロック中の他のテキストラインから分離して、以下に詳細に説明するように分離したラインごとに順々に処理を行う。ステップS207では、ライン中の各文字をライン中の他の文字から切り出し、以下に詳細に説明するように個々の文字に対して文字認識処理を行う。

【0034】ステップS208では、文字ごとに各文字の認識処理が行われ、既知の手法で文字辞書19に蓄えられている標準文字と各文字との比較処理が行われる。文字辞書19中の索引は一般には1つの文字に対応するが、分離が難しいような文字列（例えば、「f i」）や分離されやすい1文字（例えば「j」）に対しては他の索引が用意されることもある。すなわち、分離が難しいような接触文字ペアを辞書の索引とすることも、分離されやすいような1文字中の部分を辞書の索引とすることもある。比較処理に基づいて文字画像に対する識別子が選択され、選択された文字識別子はRAM20に蓄えられたり、出力装置21に出力される。また、識別された文字をプリンタ/ディスプレイ22に表示することもできる。

【0035】ステップS209では、テキストブロックに対する処理が終了したかどうか調べられる。処理が終了していなければ、処理をステップS206（あるいは適宜ステップS207）に戻し、ライン分割処理、文字切り出し処理及び文字認識処理とを繰り返す。ブロック処理が終了していれば、処理をステップS210に進め、1ページ分の処理を終了したかどうかを調べる。1ページ分の処理が終了していなければ、処理をス

テップS205に戻し、ページ中の次ブロックを選択して処理を続ける。

【0036】1ページ分の処理が終了していれば、処理をステップS210からステップS211に進め、後処理を実行する。後処理では、文脈チェックやスペルチェックなどの処理を行い、文字認識ステップS208で認識された文字識別子を、ステップS208のような個々の文字ごとではなく文字の周囲の文脈に基づいて、全体的な見地から修正する処理を行う。ステップS212では、ブロック選択ステップS203で定義された階層的木構造に基づいて画像データが再生される。ページ再生では画像データが適切な順序で配置される。例えば、脚注はメインテキストから分離され、段は他の段と混ざることなく順番に配置され、グラフィックスや線画はページ中の認識文字テキストの適当な位置に挿入される。このようなグラフィックス画像や線画の見だしも図に隣接して挿入される。この際、他の規則を用いることもでき、例えばページの物理的再生ではなく、ページからテキストのみを抽出するという規則を用いることもできる。なお、このようなページ再生の規則は、装置の情報処理機能を用いてユーザが定義できる。

【0037】ステップS213では、再生されたページが出力装置21などに蓄えられる。ここでは、ROM17に蓄えられており、CPU10で実行される他の応用プログラムを用いて、スプレッドシートやワードプロセッシング処理などの情報処理を行うことができる。そして、処理された情報（あるいは適宜処理していない情報）は、ファクシミリ16やモデム14などを用いて、またはコンピュータテキストファイルを「しゃべる」音声合成装置を用いた通常の音声電話を介してなどのさまざまな手段を用いて再送される。

【0038】図3は、図2のステップS203のブロック選別処理を説明するための一般化した処理流れ図である。ここで、ブロック選別処理は文字認識システムとは別に利用可能であることに注意されたい。すなわち、画像再生装置においてブロック選別処理を用いて、ある種のブロックに対しては画像再生の第1の手法を施し、別の種類のブロックに対しては画像再生の第2の手法を施すといったことも可能である。

【0039】ブロック選別の処理速度の向上を望む場合には、ステップS300で画像データを縮小する処理を行う。画像データの縮小処理を行った場合には、ブロック選別処理は縮小画像に対して行われる。なお、この際、図2の残りの文字認識処理（ステップS204からステップS213）に影響を及ぼさないように、ブロック選別処理終了時には、縮小前の画像データに対して選択ブロックが割り当てられる。画像データ縮小処理は、 $m \times m$ 画素ブロック中の黒画素の連結性を調べながら行われる。例えば、 3×3 画素ブロック中に連結された黒画素が2個ある場合には、 3×3 画素ブロックを1つの

黒画素に縮小する。逆に、 3×3 画素ブロック中に連結された黒画素が2個ない場合には、 3×3 ブロックを1つの白画素に縮小する。

【0040】ステップS301では、画素画像を解析して連結された要素を検出し、連結要素をサイズや他の連結要素との相対的な位置に応じて分類する処理を行う。ここで、連結要素は白画素に完全に囲まれた黒画素のグループである。すなわち、連結要素は、少なくとも1つの白画素で他の黒画素グループと完全に分離されている黒画素グループである。図4を参照して以下で詳細に説明するように、ステップS301では、連結要素を検出して、連結要素から得たサイズ情報やいくつかの統計的な値に基づいて各連結要素の分類を行う。

【0041】まず、以下で詳細に説明するように、各連結要素をテキスト要素か非テキスト要素かに分類する。さらに、非テキスト部については、フレームデータであるか、中間調画像であるか、線画であるか、表あるいはテキストデータを表のように構成したものであるか、あるいは未知の要素で分類不可能であるかなどを解析処理して決定する。また、各連結要素ごとに階層的木構造を構成して、連結要素の構造的な記述データを求め、ステップS212で説明したようにデジタルデータの再生を容易にする。

【0042】ステップS302では、水平方向に近い位置関係にあり、ギャップラインマーカを横切らないような連結要素をラインとしてグループ化する。この際、ステップS301で作成した木構造を用いて、テキスト要素と非テキスト要素とが不適切にグループ化されないようにする。また、ステップS302では、段間で垂直方向に伸びるギャップや非テキスト要素の垂直方向に伸びる境界を検出して、テキスト要素を段ごとにまとめる。ここで得られた段構造は階層的木構造に組み込まれ、適宜木構造の更新が行われる。

【0043】ステップS303では、ステップS302でグループ化されたラインの内、垂直方向に近接するラインを垂直方向にグループ化して、ブロックを形成する。ここで、非テキスト要素は垂直方向にソートされ、画像ページの境界として用いられる。なお、2つの非テキスト要素の間に位置するテキストライン要素は、他のテキストライン要素とは別に処理される。また、ステップS303では、ステップS301で分類できなかった非テキスト要素が、大きなフォントサイズのタイトルであるかどうかを判別する。タイトルであると判別されると、その部分に「タイトル」属性を付与し、木構造を更新する。ここで得られたタイトルは、ステップS212に基づくページの再生の際に有用な情報となりうる。

【0044】図4A～図4Cは、画素画像データの連結要素をどのように検出して、どのようにこれらの連結要素を分類するかを示す詳細なフローチャートである。図4A～図4Cに示される処理ステップは、ROM17に

保持されているプログラムステップに基づいてCPU10で実行される。ステップS401では、輪郭追跡を行い画素画像データの連結要素を検出する。輪郭追跡処理は図5Aに示すように画像データをスキャンすることで行われる。スキャン処理は矢印Aで示されるように画像の右下部から左に行われ、画像の右側の端に達したときに上に進むように行われる。なお、この際、逆方向すなわち左上から右下にスキャンすることもできる。スキャン処理中に黒画素を検出すると、31で示すような星形パターンの順序で隣接画素を調べ、黒画素の隣接画素もまた黒画素であるかどうかを判別する。星形パターン31には8個の番号付けされたベクトルが中心点から伸びているため、このような輪郭追跡を以下「8方向」追跡と呼ぶことにする。黒画素が隣接して存在する場合には、画像の外側輪郭を追跡し終えるまで上述の処理を繰り返す。

【0045】すなわち、図5Bに示すように、矢印A方向へのスキャン処理によって、まず文字「Q」の端に対応する点32が検出される。次いで、星形パターン31に基づいて隣接画素が判別され、文字「Q」の外側輪郭が追跡される。この際、閉輪郭の内部の輪郭追跡処理は行わない。連結要素を1つ検出して、8方向追跡で輪郭の追跡を終えると、次の黒画素を検出するまでスキャン処理を続ける。すなわち、完全な黒領域である物体34の8方向追跡を行う。また、手書き単語「non-text」の非テキスト物体35を、単語「word」を構成する個々の文字のテキスト物体36中の個々の文字と同様に追跡する。図5Aに示すスキャン処理は、画素データ中のすべての連結要素を検出し、8方向追跡して輪郭を検出するまで行われる。

【0046】次いで処理をステップS402に進め、各連結要素を矩形で囲む処理を行う。具体的には、各連結要素の周りを包囲する最も小さな矩形を描く。すなわち、図5Bに示すように、物体32の周りには矩形37が、物体34の周りには矩形39が、物体35の周りには矩形40が、テキスト物体36a、36b、36c、36dの周りには矩形41a、41b、41c、41dが描かれる。

【0047】ステップS403では、各矩形要素に対して木構造中の位置が割り当てられる。ほとんどの場合、ステップS403で得られる木構造では、画素画像中の各物体は木の根に直接つながる。これは、連結要素の外側輪郭のみを追跡しており、閉輪郭の内部は追跡していないためである。すなわち、図5Cに示すように、連結要素32に対応する矩形37はページの根に直接つながる。これに対し、非テキスト物体35に対応する矩形40や、テキスト物体36a、36bに対応する矩形41a、41bなどのように、矩形が他の矩形の内部に完全に位置するような連結要素は、囲っている連結要素（この場合には要素34）の子ノードとして位置付けられ

る。また、要素34などのように少なくとも1つの子を有する連結要素は、それ自身を「主な子ノード」として位置付ける。すなわち、図5Cに示すように、要素39が「主な子ノード」として、要素39の他の子ノード40、41a、41bとともに位置付けられる。

【0048】ステップS404では、木構造の第1レベルの各連結要素をテキスト要素あるいは非テキスト要素に分類する。この分類処理は2つのステップで行われる。第1ステップでは、連結要素の矩形と所定のしきい値サイズとの比較を行う。連結要素を囲む矩形の高さが、想定される最大フォントサイズに対応する所定の第1しきい値以上であれば、あるいは連結要素を囲む矩形の幅が、ページ幅を経験的に決められた一定値で割った値以上であれば（「5」で良好な結果が得られることがわかっている）、連結要素を非テキスト要素と分類し、要素に「非テキスト」の属性を付与する。

【0049】第2ステップでは、残りすべての要素すなわち非テキストと分類されなかった要素と、残りすべての連結要素のサイズの集合に基づいて適当に求められるしきい値との比較を行う。具体的には、非テキスト要素として分類されなかったすべての矩形の平均の高さを求め、この平均高にスカラー値を掛け合わせたものを（「2」を用いると好都合である）適当に求められるしきい値とする。そして、適当に求めたしきい値以上に大きな要素はすべて非テキストであると考え、非テキスト要素として分類する。また、適当に求めたしきい値以下の小さな要素はすべてテキスト要素であると考え、このようにして要素の分類を行い、適切な属性が付与される。

【0050】なお、図4A～図4Cの残りを参照して以下で詳細に説明するように、ここで求めた分類結果は以下で改善される。木構造の第1レベルの各要素をテキスト要素あるいは非テキスト要素として分類すると、テキスト要素の主な子ノードを含むすべての子ノードをテキスト要素として分類する。これに対して、非テキスト要素の主な子ノードは非テキスト要素として分類するが、非テキスト要素の他のすべての子ノードはテキスト要素と分類する。

【0051】ステップS405において、最初の要素を選択する。要素がテキスト要素であれば（ステップS406）、ステップS407に処理を進め次の要素を選択する。非テキスト要素を選択して処理をステップS408に進めるまで、ステップS406とS407の処理を繰り返す。ステップS408では、非テキスト要素に子ノードが存在するかどうか調べられる。例えば、図5Cに示されるように、非テキスト要素39には、非テキストの主な子ノード39とテキストの子ノード40、41a、41bとが存在する。

【0052】ステップS408において子ノードが存在する場合には、処理をステップS409に進め、要素に

フィルタ処理が施され、要素が中間調（あるいはグレイ階調）要素であるかどうかを判別される。中間調フィルタ処理では要素の子ノードを調べ、要素の子ノードのうち「雑音サイズ」以下であるようなサイズのものを求める。ここで「雑音サイズ」の要素とは、高さが画像データにおいて想定される最小のフォントサイズ以下であるような要素のことである。雑音サイズ以下のサイズの子ノード数が全体の子ノード数の半分以上であれば、要素は中間調画像として判断される。そこで、ステップS410で処理をステップS411に進め、「中間調」属性を要素に付与する。

【0053】次いでステップS412で、中間調画像中のテキスト要素をチェックする。具体的には、木構造中の中間調画像の子ノードのうちテキストサイズの子ノードを修正して、テキストサイズの要素を中間調画像の子ノードとしてではなく、中間調画像と同一レベルのノードに配置する。これにより、中間調画像中のテキストサイズの要素に対しても、適切だとすれば文字認識処理を行うことができる。この処理を終えると処理をステップS407に戻し、次の要素を選択して処理を行う。

【0054】ステップS409の中間調フィルタ処理において要素が中間調画像でないと判別された場合には、処理をステップS410からステップS413に進め、さらなる処理を施すために要素の主な子ノードを選別する。そして、処理をステップS414に進める。ステップS408において非テキスト要素が子ノードをもたないと判別される場合、あるいはステップS413においてさらなる処理を施すために主な子ノードが選択されると、ステップS414で当該要素に対してフレームフィルタ処理が施される。フレームフィルタ処理は、当該要素がフレームであるかどうかを調べ、当該要素を囲むような矩形とほぼ同一の幅/高さを有する平行水平線と平行垂直線とを検出するための処理である。具体的には、画素中の行ごとに当該要素中の連結要素内部の最大長を求めて、連結要素が調べられる。

【0055】すなわち、図6Aに示されるように、非テキスト要素42には連結要素43が含まれるが、その輪郭は8方向追跡によって44で示すように追跡処理される。行「I」における連結要素内部の最大長は、輪郭の左端45aから右端45bまでの距離 x_1 となる。また、行「J」では、連結要素内部の長さとして、連結要素の境界上の点46aと46bの距離、ならびに点47aと47bの距離との2つが存在する。ここで、点46aと46bの距離の方が点47aと47bの距離よりも長いため、距離 x_1 が行「J」における連結要素内部の最大長となる。

【0056】非テキスト要素42中のn個の行ごとに距離「x」を求め、以下の不等式を満たすかどうかを調べ、非テキスト要素がフレームであるかどうかを判別する。

10

20

30

40

50

【0057】

【数1】

$$\sum_{k=1}^N \frac{(X_k - W)^2}{N} < \text{threshold} \quad (\text{しきい値})$$

ここで、 x_k は（上述のように）連結要素内部のk行目の最大長、Wは矩形要素42の幅、Nは行数、しきい値はフレームが画像データ中で傾いていてもフレーム検出が可能となるように前もって計算した値である。ここで、1度の傾き角を許容するには、しきい値として $\sin(1^\circ)$ を1倍したものに、ステップS404で計算したテキストの平均の高さをオフセットとして加えたものを用いると良い。

【0058】上記の不等式が満たされれば、要素はフレーム要素であると判断され、ステップS415からステップS416に処理を進め、「フレーム」属性を要素に付与する。ここで、複数の属性が各要素に付与されることに注意されたい。すなわち、「フレーム-表」、「フレーム-中間調」などのようにフレームの属性が付与されることもある。

【0059】ステップS416に続いて、フレーム要素中に表データあるいは表のように組織されたデータが存在するかどうかを調べる処理を行う。そこで、ステップS417において、連結要素の内部を調べて白輪郭を求める。白輪郭は、黒画素ではなく白画素に着目するという点を除けば、上のステップS401で検出した輪郭と同一のものである。すなわち、図7Aに示されるように、非テキスト要素の内部を非テキスト要素の右下から左上へ矢印Bの方向に沿ってスキャンする。このスキャン処理によって第1の白画素が検出されると、星形パターン51の順序で白画素の隣接画素がチェックされる。星形パターン51には1から4までの数が付与されたベクトルがある。このため、このような処理に基づく白輪郭追跡処理のことを以下「4方向」白画素追跡と称する。

【0060】黒画素に囲まれた白輪郭すべてを追跡し終えるまで、4方向の白画素追跡処理を繰り返す。例えば、白輪郭追跡処理によって、黒画素セグメント52、53、54、55ならびに内部に存在する56のような他の黒画素で構成される内部輪郭の画素が追跡される。このようにして白輪郭を位置付けると、矢印B方向へのスキャン処理を、非テキスト物体に囲まれたすべての白輪郭を追跡するまで繰り返す。

【0061】ステップS418では、非テキスト要素の密度を計算する。密度は連結要素中の黒画素数を計数して、黒画素数を矩形中の全画素数で割ることで計算される。ステップS419では、非テキスト要素中で検出された白輪郭数の数をチェックする。白輪郭の数が4以上であれば、非テキスト画像が表またはテキストブロックを表のように並べたものである可能性が高い。

【0062】そこで、ステップS420において、白輪郭の充填率を求める。白輪郭の充填率は、白輪郭が非テキスト画像中の領域を占める割合のことである。例えば、図7Aに示されるように、白輪郭充填率には、57や59のような斜線領域で完全に空の白領域と、60や61のような領域で中に黒画素を含むような白領域とが含まれる。この充填率が高いと、非テキスト画像が表またはテキストデータを表のように並べたものである可能性が高い。そこで、ステップS421において充填率を

チェックする。この際、高い充填率であれば、非テキスト画像は表またはテキストデータを表のように並べたものであると考えられる。

【0063】この処理の信頼性を向上させるために、白輪郭が水平ならびに垂直方向に格子状の構造をなしているかどうかを調べる。すなわち、ステップS422において、少なくとも2つの白輪郭の境界線が水平方向ならびに垂直方向に一致しないような非格子状の白輪郭を再接続する。例えば、図7Aに示されるように、白輪郭59の左境界線62及び右境界線63は、白輪郭60の左境界線64及び右境界線65と垂直方向に一致する。そのため、これらの白輪郭は格子状に配置されており、白輪郭の再接続は行わない。同様に、白輪郭59の上部境界線66及び下部境界線67は、白輪郭70の上部境界線68及び下部境界線69と水平方向に一致する。そのため、これらの白輪郭は格子状に配置されており、これらの白輪郭の再接続は行わない。

【0064】図7Aから図7Dまでは、白輪郭を再接続する場合を説明するための図である。図7Bは非テキスト要素71を示しており、これはステップS201で示した中間調画像の2値画像へのしきい値処理などによって得られるものである。非テキスト画像71には、黒領域72と白領域74、75、76、77、78、79が存在する。ここで、これらの白領域の充填率は十分高く、ステップS421において処理は再接続ステップS422に進められたものとする。

【0065】まず、図7Cに示されるように、白輪郭75の左境界線及び右境界線と、白輪郭76の左境界線及び右境界線とを比較する。すると、これらの境界線は一致しないため、白輪郭75と白輪郭76とは再接続されて図7Cの接続白輪郭76'が生成される。次いで、図7Dに示されるように、白輪郭77の上部境界線及び下部境界線と、白輪郭79の上部境界線及び下部境界線とを比較する。すると、これらの境界線は一致しないため、白輪郭77と白輪郭79とは1つの接続白輪郭77'として結合される。再接続される輪郭がなくなるまで、これらの処理を水平方向に垂直方向に繰り返す。

【0066】すなわち、上述のように、表をなす白輪郭は再接続されにくいのに対し、中間調画像や線画などの表でないような白輪郭は再接続されやすい。そこで、ステップS423において接続率を調べる。ここで、再接

続率が高ければ、あるいは再接続処理後に残った白輪郭の数が4より少なければ、処理をステップS428に進め、以下で詳述するように、非テキスト要素に「中間調画像」もしくは「線画」の属性を付与する。ステップS423において再接続率が高くなく、且つ少なくとも4つの白輪郭が残っていれば、処理をステップS424に進め、非テキスト画像に「表」の属性を付与する。

【0067】ステップS425では、新たに属性が付与された表の内部を調べ、8方向追跡処理で連結要素を検出し分類する。ステップS426では、この新たな内部の連結要素に基づいて階層的木構造を更新する処理を行う。続くステップS427では、ステップS402からステップS404で示したような手法でもって、内部連結要素をテキスト要素もしくは非テキスト要素に再分類し、適当な属性を付与する。これらの処理を終えると、処理をステップS407に戻し次の要素が選択される。

【0068】ステップS421とS423に戻って、ステップS421で充填率が高くない場合、あるいはステップS423で再接続率が高い場合には、非テキストフレーム要素は中間調画像もしくは線画である可能性が高いと考えられる。ここで、要素を中間調画像として分類するか、あるいは線画として分類するかの決定は、要素中の黒画素の平均水平ラン長、要素中の白画素の平均水平ラン長、白画素と黒画素との比率、及び密度に基づいて行われる。一般に、暗めの画像は中間調画像であると考えられ、明るめの画像は線画であると考えられる。

【0069】具体的には、白画素の平均水平ラン長がほぼゼロ（すなわち、全体が暗めあるいはまだらな画像）であって、ステップS418で求めた密度によって要素が白っぽいというよりもむしろ黒っぽい場合（すなわち、密度が1/2程度の第1しきい値より高い）には、フレーム要素は中間調画像であると判別される。密度が第1しきい値以下である場合には、要素は線画として判別される。

【0070】白画素の平均水平ラン長がほぼゼロではなく、黒画素の平均水平ラン長より長い場合には、フレーム要素は線画であると判別される。しかし、白画素の平均水平ラン長が黒画素の平均ラン長以下（すなわち、暗めの画像）である場合には、さらなる判別処理が必要となる。具体的には、黒画素数が白画素数よりもかなり少ない（すなわち、黒画素数を白画素数で割った値が約2の第2しきい値より大である）場合には、フレーム要素を中間調要素として判別する。これに対し、黒画素数を白画素数で割った値が第2しきい値以下であっても、ステップS418で求めた密度が第1しきい値より高ければ、フレーム要素を中間調画像として判別する。これ以外であれば、フレーム要素は線画として判別される。

【0071】ステップS428においてフレーム要素が線画として判別されれば、処理をステップS429に進めて「線画」属性を付与し、ステップS430ですべて

の子ノードを削除する。なお、要素が線画として判別されると、線画要素のいかなるブロックも文字認識のために選択されることはない。そして、処理をステップS407に戻し次の要素を選択する。

【0072】一方、ステップS428においてフレーム要素が線画として判別されなかった場合には、処理をステップS431に進めて「中間調」属性を付与し、ステップS432でフレーム中間調要素のテキストサイズの子ノードを削除する。ここで、テキストサイズは、ステップ404において述べたように平均の要素の高さとして求められる。また、テキストサイズより大きなすべての子ノードは、フレーム中間調要素の子ノードのままとする。これらの処理を終えると、処理をステップS407に戻し次の要素を選択する。

【0073】ステップS419に戻り、白輪郭の数が4より小さい場合には、フレーム要素は表であるとは判別されない。そこで、処理をステップS433に進め、ステップS418で求めた密度と約0.5のしきい値とを比較する。ここで、しきい値は、フレーム内のテキスト要素や線画が占める面積は画素の半分以下であることを鑑みて選択されている。

【0074】密度がしきい値より小さければ、処理をステップS434に進め、上述のようにフレーム要素の内部構造を作成する。すなわち、処理をステップ401に戻し、フレーム要素の内部構造に対する処理を行う。ステップS433において、密度が所定のしきい値以上であった場合には、処理をステップS442に進め、フレーム要素を線画、中間調画像あるいは分類不可能（すなわち、フレームが「未知」）のどれかに分類する。

【0075】ステップS415に戻り、ステップS414におけるフレームフィルタ処理で非テキスト要素中にフレームが検出されなかった場合には、処理をステップS435に進め、非テキスト要素中に線が含まれるかどうかを調べる。線は、テキスト境界を明確にするのに適した非テキスト要素である。この際、このような線に囲まれたテキストは線のそばに位置することが多いため、テキストが線に接するということもありうる。そのため、テキストが接していても接していなくても線を検出できるように線の検出処理を行う。

【0076】テキストが接していない線を検出するには、非テキスト要素の長さ方向のヒストグラムを求める。すると、図6Bに示されるように、線のヒストグラム48はほぼ均一な分布となり、ヒストグラムの高さは線幅にほぼ等しくなる。また、線幅は、非テキスト要素の幅「W」にほぼ等しい。ここで、線幅と非テキスト要素の幅との差は、画素画像を生成する際の原文書の傾き角 θ_s によるものである。そこで、非テキスト要素が線を含むかどうかを調べるために、ヒストグラム中の各セルkの高さ49を非テキスト要素の幅Wと比較する。すなわち、これらの値間の実効値誤差としきい値とを次式

のように比較する。

【0077】

【数2】

$$\sum_{k=1}^N \frac{(cell_k - W)^2}{N} < threshold \quad (\text{しきい値})$$

ここで、しきい値は、非テキスト要素中の線の傾き角 θ_s を許容できるように設定される。例えば1°の傾き角であれば、しきい値として

【0078】

【数3】

$$\sum_{k=1}^N \left\{ \frac{k \sin(1^\circ)}{N} \right\}^2$$

を用いると良好な結果が得られる。上の不等式によってテキスト要素が接していない線が検出されなかった場合には、テキスト要素が接している線を含む要素であるかどうかを調べる処理を行う。テキスト要素に接している線が非テキスト要素中に含まれるかどうかを判別するために、要素の境界に沿って線が長く伸びているかどうかを調べる。すなわち、要素を通して線が長く伸びていれば、図6Cに示すように要素を囲む矩形の境界は線にきわめて近いところに位置する。そこで、矩形の境界近くの第1番目の黒画素位置の均一性を、境界からの距離の2乗の和を求めることで判断する。すなわち、次式の不等式を満たすかどうかを調べる（図6C参照）。

【0079】

【数4】

$$\sum_{k=1}^N \frac{X_k^2}{N} < threshold \quad (\text{しきい値})$$

ここで、2乗の和が所定のしきい値より小である場合には、テキスト要素が接している線要素と判別する。この際、しきい値として、テキスト要素が接していない線の検出におけるしきい値を用いると、良好な結果が得られる。ステップS435において線を検出すると、処理をステップS436からステップS449に進め、「線」属性を非テキスト要素に付与する。そして、処理をステップS407に戻し、次の要素を選択する。

【0080】一方、ステップS435で線が検出されなかった場合には、処理をステップS436からステップS437に進め、非テキスト要素のサイズを調べる。ここで、サイズが所定のしきい値以下であると、非テキスト要素の分類が不可能となる。なお、このしきい値は最大のフォントサイズに基づいて決められるものであり、最大フォントサイズの半分の値で良好な結果が得られる。そして、処理をステップS438に進めて「未知」属性を非テキスト要素に付与し、処理をステップS407に戻して次の要素を選択する。

【0081】ステップS437においてサイズが所定の

しきい値よりも大であった場合には、処理をステップS 439、S 440、S 441に進め、ステップS 417、S 418、S 419で述べたように、非テキスト要素内部の白輪郭を追跡し、非テキスト要素の密度を求め、白輪郭の数を調べる。ステップS 441において白輪郭の数が4以下であれば処理をステップS 442に進め、要素サイズを計算し、要素が線画や中間調画像であるのに十分な大きさであるかを調べる。このサイズ計算は、非テキスト要素の高さと幅ならびに黒画素の最大ラン長に基づいて行われる。具体的には、非テキスト要素の高さや幅が最大フォントサイズ以下であれば、非テキスト要素は中間調画像や線画であるほど大きくはないと判断し、処理をステップS 443に進め「未知」属性を付与する。また、非テキスト要素の幅は最大フォントサイズより大であるが、黒画素の最大ラン長は最大フォントサイズ以下である場合にも、処理をステップS 443に進め、「未知」属性を付与する。そのあと、ステップS 407に処理を戻し、新たな要素を選択する。

【0082】ステップS 442において非テキスト要素が線画や中間調画像であるのに十分な大きさであると判断されると、ステップS 444に処理を進め、非テキスト要素が線画であるか中間調画像であるかの判別を行う。ステップS 444からステップS 448までの処理は、それぞれステップS 428からステップS 432までの処理と同一のものであり、ここでの説明は省略する。

【0083】図4A～図4C（図3中のステップS 301）を参照して以上説明したように、画素画像中のすべての連結要素を検出して分類すると、図14に示されるような木構造が得られる。図14に示されるように、木構造の根は画素画像データのページに対応する。根からの子ノードには、テキストブロックや、未知、フレーム、絵、線などからなる非テキストブロックが存在する。また、フレームの子ノードには、テキストブロック、未知の非テキストブロック、テキストブロックを有する表、絵、線などが存在する。

【0084】図10は画素画像データの典型的なページ90を示したものであり、大きなフォントサイズのテキスト91、テキスト93などを含む表92、テキストデータ94、水平線95、もう一つのタイトル96、2段落のテキストデータ97及び見だし99を備えるフレーム線画98とタイトル100で始まり、テキストデータ101、見だし103を備えるフレーム中間調画像102、テキストデータ104、水平線105、テキストデータの最終段落106へと続く第2段とが示されている。図11は、ステップS 301に基づく処理後の同一画素画像を示したものである。図11に示されているように、画素画像データ90中の連結要素が矩形ブロックで囲まれており、矩形ブロックの内部についてはステップS 415からステップS 434までのフレーム処理で

判別される。

【0085】ステップS 302では、ステップS 301で得られたすべてのテキスト要素を、木構造の位置に関わらず水平方向にグループ化する。グループ処理は各テキスト要素の密集性ならびにその近傍関係とに基づいて行われる。なお、この際、段に対応する垂直方向に伸びたギャップが検出され保持される。以下、図8を参照して、ステップS 302で行う詳細な処理について説明を加える。図8に示される処理ステップは、ROM 17に保持されるプログラムステップに基づいてCPU 10で実行される。

【0086】ステップS 801では、非テキスト要素の左端ならびに右端から垂直にギャップラインマーカを伸ばす処理を行う。すなわち、図11に示されるように、ギャップラインマーカ109aと109bとをテキストあるいは非テキスト要素（ここでは要素95）に達するまで垂直に伸ばす。また、ギャップラインマーカ109cと109dとをテキストあるいは非テキスト要素（ここでは要素95）に達するまで垂直方向に伸ばす。同様に、残りの非テキスト要素の左端ならびに右端からギャップラインマーカを垂直に伸ばす。このようなギャップラインマーカを用いることで、画素画像データにおいて段に対応するギャップ位置の判別が容易になる。

【0087】ステップS 802において、図11のテキスト要素107のようなテキスト要素は、連結によりギャップラインマーカを横切らず、且つ他のテキスト要素と接しているかあるいは他のテキスト要素から所定のしきい値内の距離にあれば、1つのテキストラインに連結される。なお、適切なしきい値として、ステップS 404で求めた平均テキスト長に経験的に得られたスカラー値を掛けたものを用いる（「1.2」を用いると良好な結果が得られる）。また、連結処理を行うに先立って、テキスト要素間の垂直ギャップを調べ、段構造を示唆するような垂直ギャップが存在するかどうかを判別する。すなわち、図11に示されるように、ギャップ108はテキスト要素のペアの間に存在するが、このギャップはテキスト画像データ中で垂直方向に数ライン程度の長さとなるため、テキスト要素が他の要素から所定のしきい値内の距離に位置しているにも関わらず、ステップS 802ではギャップが維持される。

【0088】ステップS 803では、連結ステップS 802で連結されなかったテキスト要素のペアが連結によりギャップラインマーカを横切らず、隣のラインの第3テキスト要素と共に重なるような場合に、これらのテキスト要素のペアを連結する。このようなステップにより、段構造を示すギャップではなく単にテキストライン中のランダムな空間構成に起因するようなギャップを効果的に除去することが可能となる。例えば、図11において、ステップS 802ではギャップ108の連結は行われないが、ギャップ両側のテキスト要素は1行下のラ

インの第3テキスト要素と重なり、またギャップラインマーカを横切らないため、ステップS803においてこのギャップが除去される。

【0089】ステップS804では、これらの処理結果に基づいて木構造の適切な更新が行われる。図12は、ステップS302のグループ化処理の結果を示しており、図15はステップS302のグループ化処理で修正された木構造を示している。図12に示されるように、各々が接しているテキスト要素は、ライン110のようにテキストラインにグループ化される。すなわち、木構造中に位置するテキスト要素をテキストラインに連結する処理を行うが、111などのようにテキスト要素が木構造中のフレーム表ノードの下に位置する場合にも連結処理が行われる。ここで、このようなグループ化処理は、上のステップS417からS439で求めた白輪郭境界を横切ることはなく、表中の個々の項目が1つの連続するテキストラインにグループ化されることはない。また、左段と右段との間のギャップは維持される。さらに、非テキスト要素は再連結されない。すなわち、112や113などの要素のように、互いに所定のしきい値内の距離にあっても、非テキスト要素のグループ化は行わない。

【0090】図15には、新たなグループ化処理の結果に基づいて修正した木構造が示されている。図8（図3中のステップS302）を参照して説明したようにテキスト要素のテキストラインのグループ化を行った後には、ステップS303に示されるようにテキストラインを垂直方向にグループ化してテキストブロックを生成する処理を行う。

【0091】以下、図9を参照してこの処理について詳述する。グループ化処理は、テキストライン要素の密集性や非テキスト要素の位置に基づいて行われる。例えば、間に存在する非テキストラインは境界を示すため、これを利用すれば非テキストライン両側のテキストラインを1つのテキストブロックにグループ化することを避けられる。なお、処理は、2つの連続する非テキストライン要素間のすべてのテキストラインに対して1度に行われる。また、ステップS303では、テキスト要素を非テキスト要素に連結すべきか（例えば、非テキスト画像とそのテキスト見だし）、非テキスト要素を他の非テキスト要素に連結すべきか（例えば、中間調画像と線画）の判断も行う。

【0092】図9は、テキストラインのテキストブロックへのグループ化を示す詳細なフローチャートである。ステップS901においては、最大の予想フォントサイズより小ではあったが平均テキストサイズより大であったため、ステップS404で非テキスト要素に分類された非テキスト要素からタイトルブロックの形成を行う。隣り合う非テキスト要素のうち同じようなサイズの要素に対してはすべてグループ化を行い、タイトルブロック

を形成し、「タイトル」属性をこのグループに付与する。ここでグループ化されなかった残りすべての非テキスト要素に対しては「絵-テキスト」属性が付与される。また、それに応じて木構造の更新も行う。ここで得られたタイトルは、ページ再生時（ステップS212）に有用な情報となる。

【0093】ステップS902では、2つのテキストラインにまたがる非テキスト要素の位置を明確にする。このような非テキスト要素はテキストブロック間の境界となり、テキストラインを1つのテキストブロックにグループ化してしまうことを避けることができる。ステップS903では、2つのステップでテキストラインの垂直方向のグループ化を行い、テキストブロックを形成する。第1のステップでは、画素密度の垂直ヒストグラムを求めるなどして、段間のギャップを検出する。第2のステップでは、垂直方向に連続するテキストライン間の垂直距離がステップS404で求めたテキストの高さより小さければ、それぞれの段ごとにテキストラインのグループ化を行う。このステップS903の処理により、図2のライン114のような同一テキスト段落中のテキストラインを、テキストブロックにグループ化する。

【0094】ステップS904では、垂直ならびに水平方向に隣接するテキストブロックのグループ化処理を行う。ここで、これらのテキストブロックが非テキスト要素で分離されていず、またこれらのブロックを連結してもステップS903のヒストグラムから求めたギャップが保持されるときに、テキストブロックのグループ化が行われる。また、ブロック間の距離がステップS404の垂直の高さに基づいて計算される所定のしきい値より小であるときに、テキストブロックはグループ化される。このステップS904の処理により、図12中の段落115のラインと段落116のラインのテキストブロックはグループ化されるが、段落117と118のラインのテキストブロックはそれらの間に非テキスト要素119（線）を有するためグループ化が行われない。

【0095】ステップS905では、テキストブロックを非テキストブロックに連結すべきか、非テキストブロックを他の非テキストブロックに連結すべきかを判断する。ここで、テキストブロックの非テキスト-タイトルブロック、非テキスト-中間調ブロック、非テキスト-ライン接触ブロックとの連結は、以下に行われる。

(1) テキストブロックが非テキスト-タイトルブロックと水平方向に近い位置にあり、垂直方向に重なっている場合には、テキストブロックを非テキスト-タイトルブロックに連結する。

(2) テキストブロックがワードサイズのブロックより小さくて（水平、垂直方向とも）、テキストブロックの隣にワードサイズのテキストブロックが存在しないときには、テキストブロックを非テキスト-中間調画像プロ

ックの中に位置付ける。

(3) テキストブロックが非テキストライン接触ブロックと重なっていれば、ライン接触ブロックはテキストの下線である可能性が高いため、ライン接触ブロックをテキストブロックに変換する。

*

*【0096】また、以下の表にしたがって、非テキストブロックは他の非テキストブロックと連結される。

【0097】

【表1】

	中間調	線画	テキスト-絵	タイトル
中間調	テスト1	非連結	常に連結	非連結
テキスト-絵	テスト1	テスト2	テスト2	テスト3
線画	テスト1	テスト1	非連結	非連結
タイトル	非連結	非連結	非連結	テスト3

この表中のテストは以下の通りである。

テスト1：1つのブロックが完全に他のブロック内に位置すれば連結。

テスト2：絵-テキストの幅がワードサイズのブロックの幅より小さければ連結。

テスト3：ブロックが近接していれば連結。

【0098】ステップS906では、適切な属性が付与され、上述の処理結果に基づいて木構造の更新が行われる。図13は図9の処理を行った結果のブロック構造であり、図16はその木構造の例である。図13において、ブロックには、タイトルブロック120、テキストブロック121及び絵データ122が含まれる。また、フォームデータも含まれ、123は表構成のデータを有するフレーム要素を示しており、124はテキスト要素125を有するフレーム要素を示している。なお、非テキストライン画像127は図13に示されるさまざまな要素を分離している。

【0099】図3から図16を参照しながら説明したブロック選択処理を終えると、上述のように文字認識処理は図2のステップS204に処理を進める。すなわち、階層的木構造中の第1ブロックを選択して認識処理を行う。このブロックがテキストブロックでない場合には、処理をステップS204からステップS205に進め、木構造中の次のブロックを選択する。テキストブロックが選択されるまでステップS204とS205を繰り返し、テキストブロックが選択された時点でステップS2※

※06に処理を進め、ラインの分割処理を行う。

【0100】図17は、図2のライン分割ステップS206で実行される処理ステップを詳細に示すフローチャートである。図17に示される処理ステップは、プログラムROM17に保持されるコンピュータプログラムにしたがってCPU10で実行される。ステップS1701に先立って画像縮小処理を行っても良い。しかし、ライン分割処理や文字分割処理は水平方向の空白により影響を受けやすいため、画像縮小処理を行うにあたっては分割精度に影響を与えないような注意が必要である。すなわち、水平方向と垂直方向とでそれぞれ異なる画像縮小手法を用いることが好ましい。垂直方向では画素の結合を「OR」論理で行い、垂直方向の対象画素のうち1つでも黒画素が存在すれば黒画素が出力される。すなわち、垂直方向の2：1の画像縮小処理では、2つの垂直画素のいずれかが黒画素であれば黒画素が出力される。これに対して水平方向では画素の結合を「AND」論理で行い、水平方向の対象画素すべてが黒画素であれば黒画素が出力される。すなわち、水平方向の3：1の画像縮小処理では、3つの画素がすべて黒画素のときのみ黒画素が出力される。

【0101】垂直方向に3：1の縮小を行い、水平方向に2：1の縮小を行う場合の処理例を以下に示す（「0」は白画素を「X」は黒画素を表す）。

【0102】

【表2】

原画像	垂直縮小(OR)	水平縮小(AND)
X O X O X X X O O O O O O X O O O X X X O O X O	X X X O O X X X	X O O X

画像縮小処理を終えると、この縮小画像に対してライン分割処理と文字切り出し処理とが行われる。ここで、図2の残りの認識処理（すなわち、ステップS208からステップS213まで）に影響を及ぼさないように、ラ

イン分割処理と文字切り出し処理とを終えた時点で、文字間の切り出しは縮小していないもとの画像データに対して行う。

【0103】ステップS1701では、ステップS20

4で選択されたテキストデータブロックごとに画素密度の水平投影を求める。画素密度の水平投影は、画素画像の各行ごとに黒画素数を計数して得られる。ここで、画素密度の水平投影は全体のテキストブロックに対して求めることが好ましいが、これは本質的な点ではない。すなわち、テキストブロックを複数の画素列、例えば2あるいは3列に分割して、各列ごとに画素密度の水平投影を求めることができる。もちろん、このような処理では、本発明の処理時間短縮という利点は失われる。

【0104】ステップS1702では、水平投影のうち
10 ゼロでない領域を求め、どれかが最大フォントサイズと等しい所定のしきい値より大であるか否かを調べる。最大フォントサイズ以下であれば、水平投影はページ上のラインを均一に分割していることになり、処理をステップS1703に進める。ステップS1703では、画素密度の水平投影のうち近接している領域を連結する。この処理を図18A～図18Dを用いて説明する。図18Aは典型的なテキストブロック230を示しており、文字画像のライン231と233、ならびに雑音スポット232（すなわち、文字情報ではない黒画素）が含まれる。これに対応する画素密度の水平投影は234に示される。水平投影234に示されるように、領域235は文字「1」の上の点に対応し、領域236はライン231上の残りの文字に対応し、領域237と238は雑音スポット232に対応し、領域239はライン233上の文字に対応する。これらの各ピークにより、境界が水平投影のゼロ値となる領域が定義される。ステップS1703では、同一テキストライン上に存在するピーク235と236のような近接領域は連結され、テキストラインとは関係のないピーク237と238のような近接
20 領域は連結されないことが望まれる。

【0105】ステップS1703で述べたように、近接領域を連結するためには、画素密度の水平投影を投影の上部から下部まで（テキストブロックの上部から下部まで）調べる。第1の領域を検出すると、その下の次の領域の水平投影をスキャンして、2つの領域の高さを比較する。上の領域の最も高い要素の高さが下の領域の高さより小さく、2つの領域の間隔が上の領域中の最も高い要素の高さより小さければ、これら2つの領域を連結する。ここで、この近接領域を連結する処理は、スケール
40 に対して不変であることに注意されたい。すなわち、テキストの近接領域を連結する際には、テキストサイズが12ポイントであるとか8ポイントであるなどの知識は不要である。

【0106】そこで、図18Aに戻り、領域235の高さと領域236の高さとの比較を行うと、上の領域235の高さが下の領域の高さより小さいことがわかる。また、2つの領域間のギャップは領域235の高さより小さいと判断される。したがって、図18Bに示されるように、領域235と236とは1つの領域236'とし
50 て連結される。

て連結される。

【0107】画素密度の水平投影を下方向に調べると、領域237が検出される。この場合、領域236'の高さが領域237の高さ以上であるため、これら2つの領域は連結されない。さらに、画素密度の水平投影を下方向に調べると、領域237の高さが領域238の高さ以下であり、また領域間のギャップは領域237の高さ以下であることがわかる。そこで、図18Cに示されるように、領域237と238とは1つの領域238'として連結される。画素密度の水平投影をさらに下方向に調べると、領域239が検出される。この場合、領域238'の高さは領域239の高さより低いが、領域間のギャップは領域238'を構成している領域237と238とのどちらかの高さより大きいことがわかる。そこで、これら2つの領域は連結されない。

【0108】また、近接領域の連結処理のあとに各領域の高さを調べて各領域の高さがラインの最小の高さに対応するしきい値より大であるかどうかを判別することもできる。この際、しきい値は、これまでに検出された領域の高さの平均として適当に設定される。領域がしきい値以下の高さの場合には、領域を画素データ中の雑音スポットに起因するものと判断し削除することができる。したがって、領域238'の高さは、領域236'、238'、239の高さの平均として決められるしきい値以下であるため、図18Dに示されるように領域238'は削除される。

【0109】これらの処理を終えると、図17に戻って処理をステップS1704に進め、領域を個々のラインセグメントに分割し、図2のステップ207で示される文字切り出しに処理を進める。ステップS1702において、ステップS1701で処理された領域が大きすぎる領域であった場合には、テキストラインが傾いていると判断される。例えば、図19Aに示されているように、テキストブロック240には複数の傾いたテキストライン241が含まれる。ステップS1702に基づく処理では、244に代表テキストを示したようにテキストラインが水平方向に相互に重なりあってしまうため、242のような画素密度の水平投影が得られる。

【0110】そこで、処理をステップS1705に進め、テキストブロックを段に分割する。図19Bに示されるように、テキストブロックの段数を2倍にする。すなわち、テキストブロック240を2つの段に分割する。この際、少なくとも1つの共通の画素が重なり合うように段を分割することが好ましい。また、テキストブロックを2ブロック以上、例えば3あるいは4ブロックに分割することも可能である。

【0111】ステップS1706では図19Bの247や249のように各段ごとに画素密度の水平投影を求め、ステップS1707で領域が大きすぎるかどうかを
50 再び調べる。領域が大きすぎる場合には、処理をステッ

ブS1708に進め、段数を再び増加させる。例えば、図19Cに示されているように、段数を更に2倍にする。また、ステップS1709において、段の幅が最低限度より大であるかどうかを確認する。この最低限度は、これ以上段数を増加させると適切なライン分割が不可能になる点を示している。好適な実施例では、最低限度は16画素幅である。ステップS1709において最低限度に達していると判断されると、ライン分割は不可能であると表示して処理を終了する。一方、最低限度まで達していない場合には、処理をステップS1706に戻し、新たな段ごとに水平投影を再び計算する。ステップS1707において領域が大きすぎないと判断された場合には、処理をステップS1710に進める。すなわち、図19Cに示されるように、ラインセグメント以下の幅の領域が割り出された。そこで、処理をステップS1710に進め、ステップS1703で説明したように近接領域を連結する。そして、ステップS1711において、各段ごとに単一ラインセグメントに対応する領域を割り出す。

【0112】すなわち、図19Cにおいて、単一ラインセグメントに対応する領域250、251、252、253を検出する。そして、領域が異なる段間で接触しており単一ラインセグメントを構成するかどうかを判別するために、各段を上から下に調べ各段の第1領域を検出する。そこで、その領域に接触している領域を調べ、図19Dに示されているように2つの距離を求める。この2つの距離、(1)2つの領域を合せたときの全長距離Aと、(2)2つの領域の共有領域すなわち2つの領域の交差領域の距離Bとの2つである。そして、比A/Bを求め、2つの領域が多くの部分で重なっているかを確かめるためにしきい値と比較する(しきい値として5を用いると良好な結果が得られる)。比A/Bがしきい値より小であれば、2つのブロックは多くの部分で重複していることになり、ブロックは単一ラインセグメントを構成していると考えられる。そこで、ステップS1712において、比A/Bで求められる接触領域を単一ラインセグメントとして割り出される。

【0113】ここで、比A/Bの計算と、比A/Bとしきい値との比較処理とは、スケールに不変な処理であり、ライン中のテキストサイズに関わらず、重なった接触領域は単一ラインセグメントとして割り出される。このようなスケール不変の性質は、ラインセグメント中のテキストサイズが既知である必要がないという点で望ましいものである。

【0114】図17で述べたラインの分割処理を終えると、図2のステップS207で説明し、詳細が図20に示される文字分割(文字切り出し)に処理を進める。図20に示されるように、文字切り出しは多階層の処理で行われ、各階層は徐々に複雑な文字切り出し処理を行う。すなわち、文字切り出し処理は3つの処理に分類さ

れる。相互に接触ならびに重なっていない文字間の切り出し、相互に接触してはいないが重なっている文字間の切り出し、接触している文字間の切り出しの3つである。例えば、図34Bに示されるように、文字「S」と「a」は接触ならびに重なっていないため第1の処理に分類される。一方、文字「f」と「y」は接触してはいないが重なっているため第2の処理に分類される。さらに、文字「t」と「l」は相互に接触しているため第3の処理に分類される。

【0115】図20に示されるように、各階層は3つの処理のうちの1つを実行するように構成される。すなわち、階層1(261)では接触ならびに重なっていない文字間の切り出しが行われる。階層1に続いて、当該テキストの性質及び特徴に関する知識の有無に応じて処理を進める。テキストが単一スペースのテキストであれば、すなわち文字が垂直で文字間が均一であれば(例えば「クーリエ」フォント)、処理を階層2(262)に進め、接触している文字の切り出しを行う。単一スペースの文字であっても、コピーやファクシミリ転送の繰り返しに起因する画像の劣化のため、文字が相互に接触することがある。そして、文字認識処理263に処理を進め、図2のステップS209に進む。

【0116】一方、テキストブロックの性質及び特徴に関して何の情報もない場合、あるいはテキストが単一スペースのテキストでなければ、処理を階層2(264)に進め、接触はしていないが重なっている文字間の切り出しを行う。そして、階層1と階層2との切り出しで得られた文字に対して265で認識処理を行う。認識処理265で認識されなかったすべての文字に対しては、階層3(266)の処理を行う。すなわち、文字が認識されない理由として、文字の切り出しが完全に行われないため実際には認識不能な文字は接触している複数の文字と考えられる。そこで、階層3では接触している文字の切り出しが行われる。階層3で文字を切り出すと、267で文字認識処理を行う。認識処理が成功した場合には、処理を図2のステップS209に戻す。一方、認識処理が再び失敗した場合には、階層3での切り出しが不適切であったと判断する。そこで、認識不能であった切り出しを再接続し、再び階層3の切り出し処理と認識処理とを行う。このような処理を、文字の切り出しが不可能になるまで繰り返し行う。

【0117】「単一スペース(mono-spaced)」262か「全スペース(all-spacing)」264かのどちらに処理を進めるかは、オペレータからの入力によって選択される。オペレータからの入力がない場合には、デフォルトとして処理を「全スペース」264に進める。というのは、この階層は単一スペースのみならず非単一スペース文字にも適用できるためである。

【0118】図21から28までは階層1から階層3までの処理を説明する図であり、図30と図31は269

で示される再接続処理を説明する図である。図21は非接触で重なっていない文字を切り出す階層1の切り出し処理を説明するフローチャートである。階層1では、2つの文字間の白画素すなわち空白を検出することにより、非接触で重なっていない文字の切り出しを行う。

【0119】具体的には、図21のステップS2101に示されるように、空白でない画素すなわち黒画素が検出されるまでラインセグメント中をとびとびに検索し、文字間の白スペースを検出する。ここで、とびとびの検索とは、ラインセグメントのすべての画素を検索するのではなく、図22に示されるように、ラインセグメントの1つの段中の一部の画素271のみを検索するものである。なお、ラインセグメントの段中の全画素の1/3のみ、すなわち3画素ごとの検索で十分であるとの結果が得られている。画素271中で空白でない画素すなわち黒画素が検出されなかった場合には、272で示すように数列の画素をとびとびとして、例えば3画素ごとに新たな列で黒画素を検索する。ここで、画素を3列とびとびとして、非接触で重なっていない文字の検出性能は低下しないとの結果が得られている。このようなとびとびの検索を、図22の画素274のような最初の黒画素が検出されるまで繰り返す。

【0120】最初の黒画素が検出されるとステップS2102に処理を進め、ラインセグメントを後向きに検索し、完全に空白な列を検出する。ここでの検索は、ステップS2101の検索とは異なり、各列ごとにすべての画素を検索して完全に空白な列を検出するものである。すなわち、図22に示されるように、完全に空白な列276が検出されるまで後向きステップ275が実行される。

【0121】完全に空白な列が検出されると、処理をステップS2103に進め、完全に空白な列が検出されるまで画素274の列から前向きに検索される。ステップS2102と同様に、ここでの前向き検索も各列ごとにすべての画素を検索するもので、277で示すように前向きに検索される。図22の278で示されるような完全に空白な列が検出されるまで前向き検索が行われる。

【0122】ステップS2103で完全に空白な列が検出されると、処理をステップS2104に進め、空白列276と278とで文字を切り出す。その後、処理をステップS2101に戻し、空白でないすなわち黒画素が再び検出されるまで、ラインセグメントでのとびとびの検索を再び行う。ラインセグメント全体で階層1の処理を終えると、テキストが単一スペース（クーリエフォントなど）であるか、あるいはテキストのスペースが未知もしくは単一スペース以外（比例フォントなど）であるかに応じて、図20の261あるいは264の階層2の処理に進める。テキストが単一スペースであれば、261の階層2の切り出しに処理を進める。

【0123】図23は単一スペーステキストのための階

層2の処理の流れを示すフローチャートである。なお、図23に示される処理ステップは、ROM17に保持されるプログラムステップに基づいてCPU10で実行される。階層2の処理に先立って、文字セグメントの幅を調べて、過小サイズの文字セグメントを割り出す。ここで、文字セグメントの幅が平均文字幅の半分以下のときに過小サイズであると判断する。過小サイズの文字セグメントが隣り合って存在する場合には、階層1の処理で1つの文字を半分ずつ2つに切り出してしまった可能性が高いため、この過小サイズ文字のペアを連結する。

【0124】ステップS2301では、各文字ブロックの幅をすべての文字ブロックの平均幅と比較して、階層1で切り出された文字ブロックのうち過大サイズのものを割り出す。ここで、各文字ブロックの幅とすべての文字ブロックの平均幅との比較は、文字が単一スペースであって各文字がほぼ同一の幅を有することが既知であるため、過大サイズの文字ブロックの検出には有効な処理となる。文字ブロックの幅（「W」）が以下の式を満たすと、文字ブロックが過大サイズであると判断される。

【0125】

$$[数5] W > (1+c) * W_{ave}$$

ここで、cは定数、 W_{ave} はラインセグメント中のすべての文字ブロックの平均幅である。なお、この判別処理はスケールに不変であることに注意されたい。定数cは、以下のように単一スペースフォントの統計的性質に基づいて決定される。クーリエアルファベットなどの単一スペースアルファベットの各文字は、単一スペースを有しており、各スペースは文字が存在する部分 α_i と文字の周りの空白スペースの部分 β_i とからなる。

【0126】例えば、図24の文字「e」に示されるように、文字「e」が存在するスペースは中心領域 α_i と周りの空白スペース β_i とからなる。ここで、iは「e」に対応する番号であり、すべてのiに対して $\alpha_i + \beta_i = 1$ である。 α_i と β_i についてはアルファベット中の各文字、すなわち英字、数字、記号などごとに求めることができ、 α_i と β_i の平均値（それぞれ α と β ）ならびに標準偏差（それぞれ σ_α と σ_β ）を計算できる。そこで、定数cは以下の式で求める。

【0127】

$$[数6] c = \sigma_\alpha / \alpha$$

図1の装置で用いるクーリエ文字セットでは、 $\alpha = 25/35$ 、 $\sigma_\alpha = 10/35$ となるため $c = 0.4$ となる。ステップS2301で過大サイズの文字ブロックが割り出されると、ステップS2302に処理を進め、過大サイズブロック中に含まれるおおよその文字数を算出し、おおよその文字境界を求める。具体的には、図24において、ブロック280の幅Wは、すべてのブロック280から283の平均の幅に $(1+c)$ を乗算して計算されるしきい値より大であるため、ブロック280は過大サイズの文字ブロックと判断される。そして、幅W

を α で割った値を最も近い整数値に丸めて、過大サイズブロック280中のおおよその文字数を算出する。

【0128】

【数7】文字数 $N = [W/\alpha]$ (最も近い整数)

また、ここで得られたブロック中のおおよその文字数「N」に基づいて、過大サイズのブロックを単一に分割して、おおよその文字境界を求める。ステップS2303では、ブロック中の画素の垂直投影特性284を、おおよその文字境界の近傍285で求める。ここで、垂直投影特性284を求める近傍は、距離 σ に基づいて決

められる。すなわち、図24に示されるように、おおよその文字境界の近傍 $\pm\sigma$ で垂直投影特性284を求める。【0129】ステップS2304では、各垂直投影特性284中での最小位置286を割り出し、この最小位置284で文字の切り出しを行う。図20の261の階層2の処理を終えると、文字認識263、そしてさらに図2のステップS209に処理を進める。ラインセグメント中の文字のスペースが未知あるいは単一スペースでない場合には、文字がラインセグメント中で単一のスペースを有しているとは限らない。そこで、図20の264の階層2の処理に進み、非接触であるが重なっている文字間の切り出しを行う。図25と図26とはこの処理を説明する図である。

【0130】ステップS2501において、階層1で切り出された各文字を分析して、文字ブロック中の各画像の輪郭のアウトラインを追跡する処理を行う。すなわち、図26Aに示されるように、文字「fy」を有する文字ブロックは非接触だが重なっている文字「f」と「y」を含み、これらは重なっているため階層1の処理では切り出されない。そこで、図26Bに示されるように、まず、この文字ブロックをブロックの右下から左方向ならびに上方向に調べて黒画素を検出する。黒画素を検出すると、図26Cの287のように黒画素と黒画素とを接続したものである輪郭の追跡が行われる。第1の文字に対して全体の輪郭追跡を終えると、288のように文字ブロック中のすべての黒画素の輪郭を追跡するまで、スキャン処理を続ける。このような処理により、各々が分離した非接触の文字が得られ、図26Dに示されるようにこれらの文字が文字ブロックから切り出される。

【0131】階層2の処理では、非接触だが重なっている文字の切り出しとともに、複数のストロークからなる単一文字、例えば「i」、「j」、「:」、「;」、「!」、「=」、「%」をも分離してしまう。そこで、ステップS2502でこのような文字の再接続を行う。図27はこの処理の詳細なフローチャートである。図27に示す再接続処理の対象となるのは階層2の処理で切り出された文字のみであり、特定の条件を満たすときのみ再接続が行われる。具体的には、ブロックが重なる、

すなわち左側文字の最も右側の画素が右側文字の最も左側の画素の上あるいは下に位置するような場合にのみ、ブロックの再接続が行われる。

【0132】そこで、ステップS2701でブロックが重なっているかどうかを調べる。ブロックが重複していなければ、再接続処理は不必要であり(ステップS2702)、再接続処理を終了する。一方、ブロックが重複していれば、ステップS2703に処理を進め、ブロックが垂直方向に分離しているかどうかを調べる。ここで、ブロックが垂直方向に分離していると、「i」、「j」、「:」、「;」、「!」、「=」などの複数ストローク文字が階層2の処理で切り出されてしまった可能性があるため、これらの文字であるかどうかを調べる。

【0133】これらのブロックは垂直方向に分離されているため、高さがH1の上部ブロックと高さがH2の下部ブロックとを含む。これらの高さの計算はステップS2704で行われ、ステップS2705でH2が(2×H1)より大であれば文字「i」あるいは「j」が分離したものである可能性が高い。そこで、文字の重複度を計算する(ステップS2706)。すなわち、2つの部分に隣接する最も右側の4画素列を平均して、この平均位置の差を計算する。(全体位置ではなく最も右側の位置を平均したのは、文字「i」や「j」の上のドットが「i」や「j」の中心ではなく上部のセリフの右側に位置するためである)。この際、スケール不変の処理とするために、平均位置の差が本体のうちの小さい幅に定数を乗じたものより小であれば、ブロックを再接続するものとする。ここで、定数は、分離の程度が予測できないような劣化画像でも再接続できるように選択され、本実施例では「9/8」としている。

【0134】一方、ステップS2705でH2が(2×H1)以下、すなわち下部ブロックの高さが上部本体の2倍以下である場合には、「:」、「;」、「!」、「=」などの文字が切り出された可能性が高い。そこで、ステップS2707に処理を進め、この可能性を調べる。すなわち、各本体中の隣接する4画素列の平均中心値を求め、これらの中心値の差を得る。この際、スケール不変の処理とするために、中心値の差が2つの本体の幅のうちの小さい方に定数を乗じたものより小であれば、上述の文字の1つである可能性が高いためブロックを再接続するものとする。なお、上述のように、定数として「9/8」を用いると良好な結果が得られている。

【0135】ステップS2703でブロックが垂直方向に分離されない場合(すなわち、2つのブロック間に水平方向に伸びるギャップが存在しない場合)には、文字はタイプ1のパーセント記号(「タイプ1」)、タイプ2のパーセント記号(「タイプ2」)、タイプ3のパーセント記号(「タイプ3」)のどれかである可能性が高い。そこで、ステップS2708で以下のように順々に

各タイプごとにチェックする。なお、変数は以下の通りである。

【0136】W1:第1文字の幅(左から右)

W2:第2文字の幅

H1:第1文字の高さ(上から下)

H2:第2文字の高さ

L1:第1文字の左端画素の列

R1:第1文字の右端画素の列+1画素

L2:第2文字の左端画素の列

R2:第2文字の右端画素の列+1画素

注意:L1は常にL2以下である。

【0137】まず、タイプ1のパーセント記号をチェックする。以下の2つの条件が満たされるときにタイプ1のパーセント記号であると判断され、ブロックが連結される。

1) $0.24 < \min(W1, W2) / \max(W1, W2) < 0.77$

これはドット幅とライン幅との比の条件である。

【0138】2) $[\min(R1, R2) - \max(L1, L2)] / \min(W1, W2) > 0.76$

これはブロックが水平方向に大部分重なっているという条件である。次いで、タイプ2のパーセント記号をチェックする。以下の4つの条件が満たされるときにタイプ2のパーセント記号であると判断され、ブロックが連結される。

【0139】1) $(0.25) L2 < R1 - L2$

これはブロックが水平方向に十分重なっているという条件である。

2) $0.50 < W1/W2 < 1.10$

これはドット幅とライン幅の適切な比率の条件である。

3) $0.43 < (H1/H2) < 0.70$

これはドット高とライン高の適切な比率の条件である。

【0140】4) $(1/m) > 0.37$

ここで、mはパーセント記号の「斜線」部上のP1とP2とを結ぶ線の傾きである。なお、P1とP2とは以下の手法で求められる。

P1:P1は第2文字の上部からD行目の行で、第2文字のプリントテキストを含む左端の画素の位置である。ここで、変数DはD=(0.1)W2である。

【0141】P2:P2は第2文字の下部からD行目の行で、第2文字のプリントテキストを含む左端の画素の位置である。

さらに、タイプ3のパーセント記号をチェックする。以下の条件が満たされるときにタイプ3のパーセント記号であると判断され、ブロックが連結される。

1) $(0.25) L1 < R2 - L1$

これはブロックが水平方向に十分重なっているという条件である。

【0142】2) $0.50 < W2/W1 < 1.10$

これはドット幅とライン幅の適切な比率の条件である。

3) $0.43 < (H2/H1) < 0.70$

これはドット高とライン高の適切な比率の条件である。

4) $(1/m) > 0.37$

ここで、mはパーセント記号の「斜線」部上上のP1とP2とを結ぶ線の傾きである。なお、P1とP2とは以下の手法で求められる。

【0143】P1:P1は第1文字の上部からD行目の行で、第2文字のプリントテキストを含む右端の画素の位置である。ここで、変数DはD=(0.1)W2である。

P2:P2は第1文字の下部からD行目の行で、第2文字のプリントテキストを含む右端の画素の位置である。

図20の264で説明した(また、図23から図27で詳細に説明した)階層2の切り出し処理と再接続処理を終えると、265の認識処理を切り出された文字に対して行う。階層1と階層2の切り出しにおいてラインセグメント中のほとんどの文字は適切に切り出されているため、265の認識処理では階層1と階層2で切り出された文字のほとんどを認識することができる。これに対し、265で認識不能であった文字は、その文字ブロックに接触文字が含まれている可能性が高い。そこで、このような認識不能の文字ブロックに対して、266の階層3の切り出し処理を行い、接触文字を切り出す。

【0144】図28は階層3の切り出し処理を示すフローチャートであり、図29Aから図29Dは接触文字を階層3で切り出す処理を説明する図である。図28に示される処理ステップはROM17に保持されており、CPU10で実行される。一般に階層3の切り出し処理は、文字ブロックを斜めに切り出して行われる。斜めの切り出し線の傾きと位置とは、ブロック中の画素密度の垂直投影特性を求め、垂直投影特性中で最も深い谷の側面の傾きを求めることで算出される。そこで、再び画素密度の垂直方向以外の投影を行う。すなわち、垂直投影特性中の谷側面の傾きに対応する回転角度方向に画素密度の投影を行う。こうして得られた複数の密度投影の中での最小点を検出し、最小点が得られた角度と位置で切り出しを行う。以下、この処理について詳細な説明を行う。

【0145】ステップS2801において、画素密度の垂直投影特性を求める。例えば、図29Aに示されるように、接触文字「t1」に対して垂直投影特性を求める。ステップS2802では、垂直投影特性中の第1の谷を検出する。垂直投影特性はデジタルであるため(すなわち、離散的な画素数の和であるため)、滑らかではなく、谷は垂直投影特性中の最小値が第1の低い値より下であって、第2の高い値より大きい極大点で両側が囲まれた点であることで見付け出される。したがって、図29Aに示されるように、垂直投影特性を調べて、高いしきい値292より高い上の点で囲まれてい

るような低いしきい値291より下の点が存在するかどうかを判別する。このような条件を満たす点が検出されれば、処理をステップS2803に進める。このような条件を満たす点が検出されなければ、以下に述べるようにしきい値の変更を行う。

【0146】まず、低いしきい値291を垂直投影特性の最大値の10%とし、高いしきい値292を垂直投影特性の最大値の20%とする。ここで、高いしきい値と低いしきい値に関する条件を満たす点が検出されない場合には、高いしきい値と低いしきい値との双方とも垂直投影特性の最大値の2%だけ増加させる。図29Aでは、しきい値291と292の条件を満たす点は検出されない。そこで、しきい値を図29Bに示すように増加させて、低いしきい値291以下であって、高いしきい値292より高い点295と296が両側に存在するような点294を検出する。点294を検出すると、ステップS2803に処理を進め、点294を囲む谷側面の傾きを算出する。谷の右側面の傾きは点294と295とを結ぶ線の傾きであり θ_1 で示される。同様に、谷の左側面の傾きは点294と296とを結ぶ線の傾きであり θ_2 で示される。

【0147】そこで、ステップS2804に処理を進め、角度 θ_1 と θ_2 および角度 θ_1 と θ_2 の近傍で回転させて投影特性を求める。すなわち、回転投影特性を角度 θ_1 、 $\theta_1 \pm 3^\circ$ 、 $\theta_1 \pm 6^\circ$ 、 θ_2 、 $\theta_2 \pm 3^\circ$ 、 $\theta_2 \pm 6^\circ$ で求める。この回転投影特性は文字ブロック中の画素を三角関数変換することで求められる。より簡易な手法として、各回転角度(最も近い角度に近似)ごとに画素位置が示される表を用意して、テーブルルックアップで回転投影特性を求めることもできる。なお、回転投影特性の各点は、この画素位置の和として求められる。

【0148】図29Cの297は回転投影特性の典型例を示したものである。回転投影特性297上の各点は、ここでは θ_1 の回転方向の画素数の和として求められる。上述のように、ここでの和は、文字ブロック中の画像を三角関数変換することで求めることもできるし、各回転角度ごとに用意された表を参照して簡易に求めることもできる。

【0149】図29CとDの点線で示されるようなすべての回転投影特性を求めると、ステップS2805に処理を進め、各回転投影特性(10個すべて)ならびにステップS2801で求めた垂直投影特性とを比較して、すべての投影特性の中で最小点を検出する。最小点となる投影特性の角度が切り出し角度に対応する。すなわち、図29Cに示されるように、点299が11個の投影特性の中で最小点となれば、文字ブロックの切り出しが角度 θ_1 で最小点299の位置で行われる(ステップS2806)。

【0150】階層3の切り出し処理を終えると、切り出

された文字に対して認識処理267を行う。ここで、階層3で切り出された2つの文字ブロックとも認識が行われたならば、図2のステップS209に処理を進める。しかし、ここでもなお認識不能なブロックが残る可能性があり、不適切な切り出しが行われた可能性と切り出された文字ブロックを再接続すべきであるという点を考慮する必要がある。この処理は再接続ブロック269で行われ、図30に詳細に示されている。

【0151】ステップS3001では、階層3での切り出し部分の双方に対して267の認識処理を試みる。ステップS3002で2つの要素の認識が可能であると判断されると、上述の図2のステップS209に処理を進める。一方、2つの切り出し要素とも認識不能である場合には、ステップS3003において少なくとも1つの要素が認識可能であるかどうかを判別する。2つの要素とも認識不能である場合にはステップS3004に進み、各要素に対して階層3の切り出し処理とステップS3001などの処理とを行う。

【0152】一方、ステップS3003において少なくとも1つの要素が認識可能であると判断されると、ステップS3005に処理を進め、認識不能の要素に対してさらに階層3の切り出し処理を行う。そして、ステップS3006で新たに切り出された要素が認識可能であると判断されれば、認識不能な要素はなくなり図2のステップS209に処理を進める。これに対して、新たに切り出された要素がともに認識不能であるとステップS3006で判断されると、ステップS3007に処理を進め、不適切な切り出しブロックを再接続することを考える。

【0153】図31は、階層3の処理において不適切な切り出しが行われる可能性を説明する図である。図31Aはイタリック文字「hm」がかなり劣化している例を示したものである。図20の266の階層3の処理では第1の切り出しを301で行うため、文字「h」の垂直部とループ部とが分離してしまう。そして、切り出された各要素に対して267で認識処理を行い、第1のブロックは「l」として認識され、第2のブロックは認識不能と判断されたとする。

【0154】このような場合、処理はステップS3005に進み、認識不能の要素303に対して図31Bのようにさらなる階層3の切り出し処理が行われる。階層3の処理により、304でさらなる切り出しが行われ、切り出された要素305と306に対して267の認識処理が行われる。しかし、要素305は文字として認識不能であるため、再接続処理が必要であると判断される。

【0155】そこで、ステップS3007において、認識不能の切り出し要素を、あらかじめ切り出された隣接する要素と接続する。この際、隣接する要素は認識されたものでも、認識不能のものでも良い。すると、図31Cのように、要素302を要素305と再接続して文字

「h」をほとんど含む新たな要素302'が生成される。そこで、要素302'と要素306とに対して267の認識処理を行う(ステップS3008)。

【0156】これらの処理を終え、ステップS3001に処理を進め、2つの要素ともに認識されたかどうかを調べる。この例の場合には、2つの要素とも文字「h」、「m」として認識されたため処理を終了する。一方、2つの要素とも認識不能であった場合には、上述の処理を繰り返す。

【0157】

【発明の効果】以上説明したように、本発明により、高速かつ正確に文書上の文字を認識してテキストファイルを作成する文字認識方法及び装置を提供できる。また、高速かつ正確に文書上のテキストと非テキストとを選別して、テキストブロックを割り出す方法及び装置を提供できる。

【0158】また、高速かつ正確に文書上の非テキストを分類できる方法及び装置を提供できる。また、高速かつ正確にテキストブロックからテキストラインを分割できる方法及び装置を提供できる。また、高速かつ正確に傾いた文書のテキストブロックからテキストラインを分割できる方法及び装置を提供できる。

【0159】また、高速かつ正確にテキストラインから文字を切り出す方法及び装置を提供できる。また、不適切に切り出された文字の再切り出しが可能な方法及び装置を提供できる。

【図面の簡単な説明】

【図1】本実施例の装置のブロック図である。

【図2】文字認識処理の流れを示すフローチャートである。

【図3】本実施例に基づくブロック分類と選別の処理の流れを示すフローチャートである。

【図4A】画素画像データ中の連結部位の分類処理の流れを示すフローチャートである。

【図4B】画素画像データ中の連結部位の分類処理の流れを示すフローチャートである。

【図4C】画素画像データ中の連結部位の分類処理の流れを示すフローチャートである。

【図5A】輪郭追跡を説明するための図である。

【図5B】輪郭追跡を説明するための図である。

【図5C】輪郭追跡を説明するための図である。

【図6A】非テキスト部位の分類処理を説明するための図である。

【図6B】非テキスト部位の分類処理を説明するための図である。

【図6C】非テキスト部位の分類処理を説明するための図である。

【図7A】白輪郭処理を説明するための図である。

【図7B】白輪郭処理を説明するための図である。

【図7C】白輪郭処理を説明するための図である。

【図7D】白輪郭処理を説明するための図である。

【図8】他のテキスト部位のサイズと近さに基づいてテキスト部分を水平方向に選択的に連結してテキストラインを形成する処理の流れを示すフローチャートである。

【図9】他のテキストラインのサイズと近さに基づいてテキストラインを垂直方向に選択的に連結してテキストブロックを形成する処理の流れを示すフローチャートである。

【図10】代表的な画像の画素データを示す図である。

10 【図11】ブロック分類と選別処理を説明するための図である。

【図12】ブロック分類と選別処理を説明するための図である。

【図13】ブロック分類と選別処理を説明するための図である。

【図14】図11から13にそれぞれ対応する典型的な階層の木構造である。

【図15】図11から13にそれぞれ対応する典型的な階層の木構造である。

20 【図16】図11から13にそれぞれ対応する典型的な階層の木構造である。

【図17】本実施例のライン分割処理の流れを示すフローチャートである。

【図18A】本実施例のライン分割処理を説明するための図である。

【図18B】本実施例のライン分割処理を説明するための図である。

【図18C】本実施例のライン分割処理を説明するための図である。

30 【図18D】本実施例のライン分割処理を説明するための図である。

【図19A】本実施例のライン分割処理を説明するための図である。

【図19B】本実施例のライン分割処理を説明するための図である。

【図19C】本実施例のライン分割処理を説明するための図である。

【図19D】本実施例のライン分割処理を説明するための図である。

40 【図20】本実施例の文字切り出し処理の機能ブロック図である。

【図21】図20の階層1の文字切り出し処理の流れを示すフローチャートである。

【図22】階層1の切り出し処理を説明するための図である。

【図23】図20における単一スペースモード(クエリエフォントなど)の階層2の文字切り出し処理の流れを示すフローチャートである。

50 【図24】階層2の文字切り出し処理を説明するための図である。

【図25】図20における全スペースモード（比例スペースなど）の階層2の文字切り出し処理の流れを示すフローチャートである。

【図26A】階層2の文字切り出し処理を説明するための図である。

【図26B】階層2の文字切り出し処理を説明するための図である。

【図26C】階層2の文字切り出し処理を説明するための図である。

【図26D】階層2の文字切り出し処理を説明するための図である。 10

【図27】階層2の処理で切り出された複数ストローク文字を再融合するための再融合手法を示すフローチャートである。

【図28】図20の階層3の処理の流れを示すフローチャートである。

【図29A】階層3の文字切り出し処理を説明するための図である。

【図29B】階層3の文字切り出し処理を説明するための図である。

【図29C】階層3の文字切り出し処理を説明するための図である。 20

【図29D】階層3の文字切り出し処理を説明するための図である。

【図30】図20の階層3の処理で切り出された部位の再融合処理の流れを示すフローチャートである。

【図31A】再連結処理を説明するための図である。

【図31B】再連結処理を説明するための図である。

【図31C】再連結処理を説明するための図である。

【図32】文字認識すべき文書のページの代表例である。 30

【図33A】従来のライン分割手法を説明するための図である。

【図33B】従来のライン分割手法を説明するための図である。

【図33C】従来のライン分割手法を説明するための図である。

【図34A】従来の文字切り出し手法を説明するための図である。

【図34B】従来の文字切り出し手法を説明するための図である。 40

【図35】本実施例のブロック選別プログラムのソースコードを示す図である。

【図36】本実施例のブロック選別プログラムのソースコードを示す図である。

【図37】本実施例のブロック選別プログラムのソースコードを示す図である。

【図38】本実施例のブロック選別プログラムのソースコードを示す図である。

【図39】本実施例のブロック選別プログラムのソース 50

コードを示す図である。

【図40】本実施例のブロック選別プログラムのソースコードを示す図である。

【図41】本実施例のブロック選別プログラムのソースコードを示す図である。

【図42】本実施例のブロック選別プログラムのソースコードを示す図である。

【図43】本実施例のブロック選別プログラムのソースコードを示す図である。

【図44】本実施例のブロック選別プログラムのソースコードを示す図である。

【図45】本実施例のブロック選別プログラムのソースコードを示す図である。

【図46】本実施例のブロック選別プログラムのソースコードを示す図である。

【図47】本実施例のブロック選別プログラムのソースコードを示す図である。

【図48】本実施例のブロック選別プログラムのソースコードを示す図である。

【図49】本実施例のブロック選別プログラムのソースコードを示す図である。

【図50】本実施例のブロック選別プログラムのソースコードを示す図である。

【図51】本実施例のブロック選別プログラムのソースコードを示す図である。

【図52】本実施例のブロック選別プログラムのソースコードを示す図である。

【図53】本実施例のブロック選別プログラムのソースコードを示す図である。

【図54】本実施例のブロック選別プログラムのソースコードを示す図である。

【図55】本実施例のブロック選別プログラムのソースコードを示す図である。

【図56】本実施例のブロック選別プログラムのソースコードを示す図である。

【図57】本実施例のブロック選別プログラムのソースコードを示す図である。

【図58】本実施例のブロック選別プログラムのソースコードを示す図である。

【図59】本実施例のブロック選別プログラムのソースコードを示す図である。

【図60】本実施例のブロック選別プログラムのソースコードを示す図である。

【図61】本実施例のブロック選別プログラムのソースコードを示す図である。

【図62】本実施例のブロック選別プログラムのソースコードを示す図である。

【図63】本実施例のブロック選別プログラムのソースコードを示す図である。

【図64】本実施例のブロック選別プログラムのソース

コードを示す図である。

【図 90】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 91】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 9 2】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 9 3】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 9 4】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 9 5】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 9 6】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 97】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 98】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 9 9】本実施例のブロック選別プログラムのソースコードを示す図である。

【図100】本実施例のブロック選別プログラムのソースコードを示す図である。

【図101】本実施例のブロック選別プログラムのソースコードを示す図である。

【図102】本実施例のブロック選別プログラムのソースコードを示す図である。

【図103】本実施例のブロック選別プログラムのソースコードを示す図である。

【図104】本実施例のブロック選別プログラムのソースコードを示す図である。

【図105】本実施例のブロック選別プログラムのソースコードを示す図である。

【図106】本実施例のブロック選別プログラムのソースコードを示す図である。

【図107】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 108】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 109】本実施例のブロック選別プログラムのソースコードを示す図である。

【図１１０】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 111】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 112】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 113】本実施例のブロック選別プログラムのソースコードを示す図である。

【図 1 1 4】本実施例のブロック選別プログラムのソー

70

グラムのソースコードを示す図である。

【図 232】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 233】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 234】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 235】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 236】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 237】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 238】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図239】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 240】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 241】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 2 4 2】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 2 4 3】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 244】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

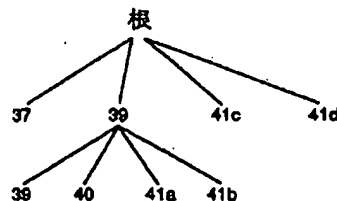
【図 245】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図246】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 247】本実施例の単一スペース（輪郭）分割プログラムのソースコードを示す図である。

【図 231】本実施例の単一スペース（輪郭）分割プロ

【图 2 6 B】



fy

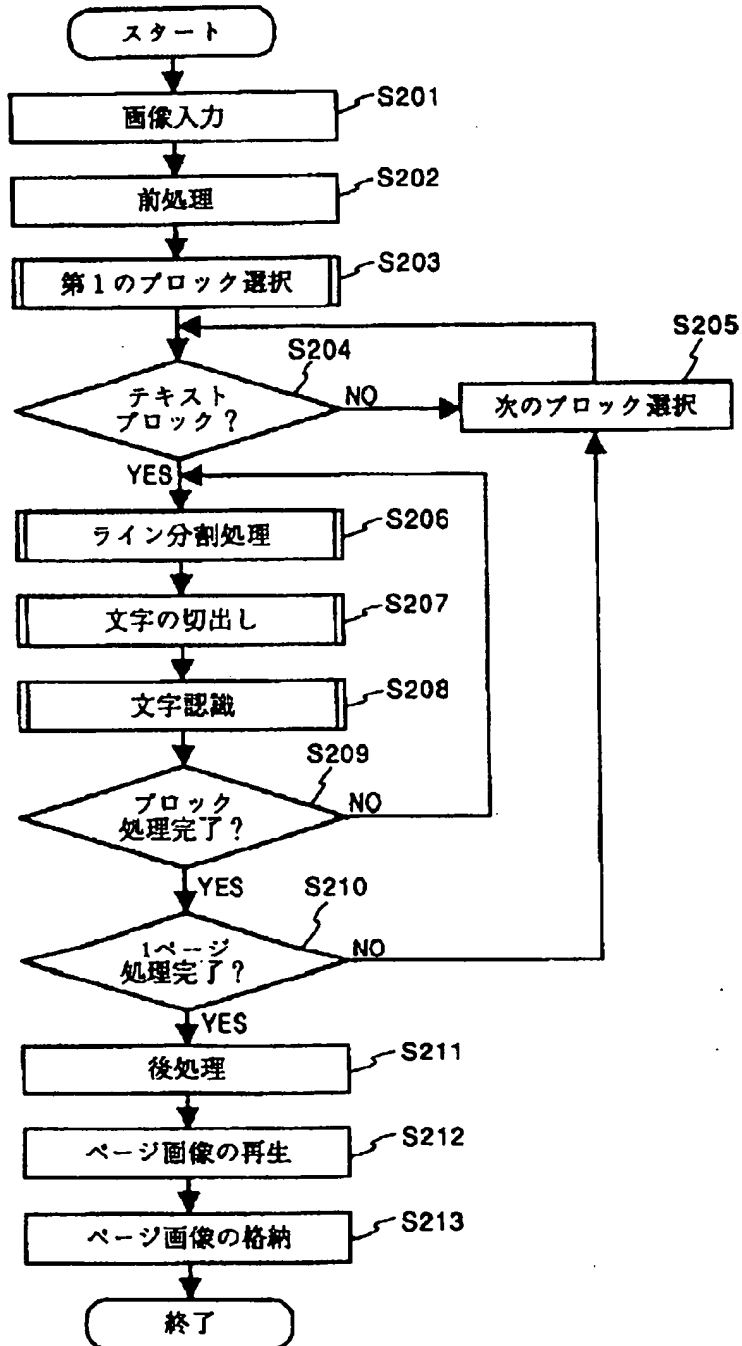
fy

```
#define FILE_LENGTH      80
#define LINE_LENGTH      256
#define FONT_NAME_LENGTH 30
```

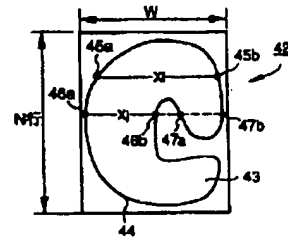
【例 36】

[illegible]

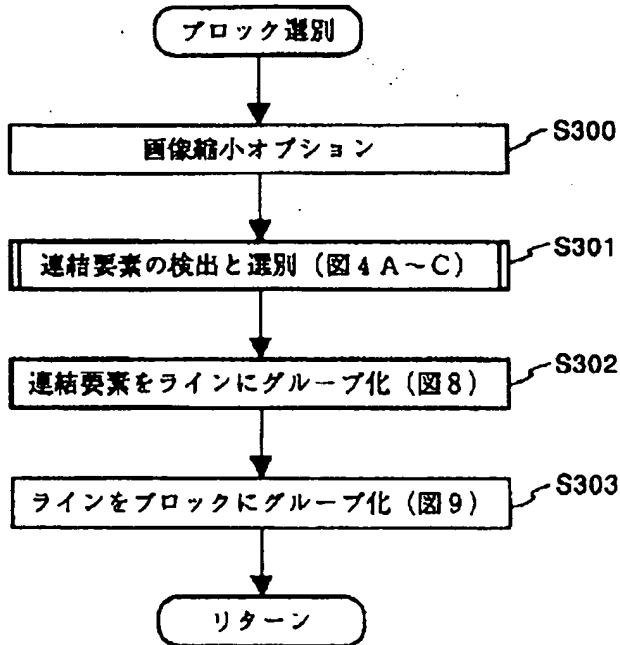
【図2】



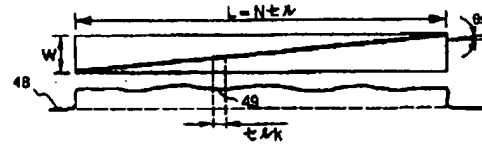
【図6A】



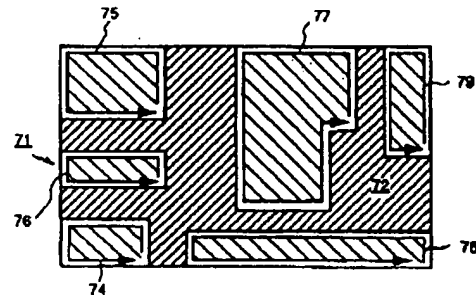
【図3】



【図6 B】



【図7 B】



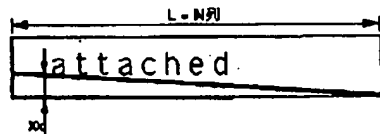
【図194】

```

free(char **M);

```

【図6 C】



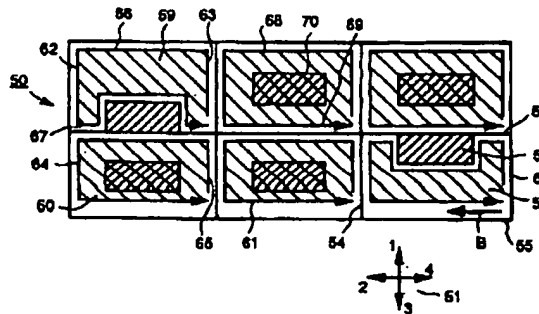
【図234】

```

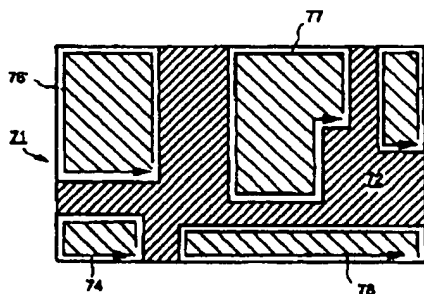
free(char **M);

```

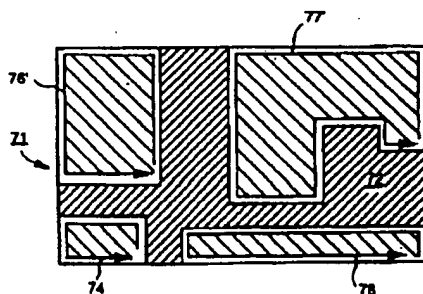
【図7 A】



【図7 C】



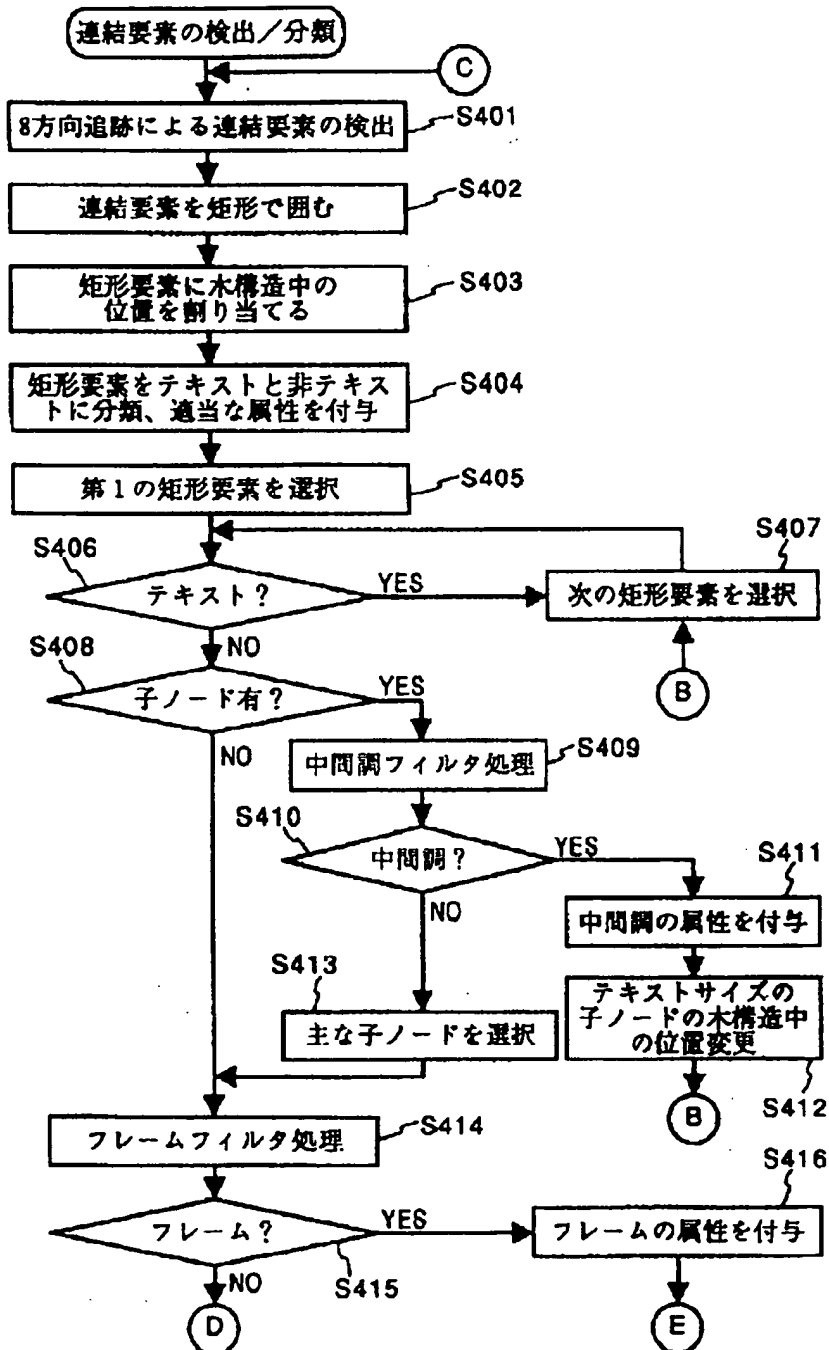
【図7 D】



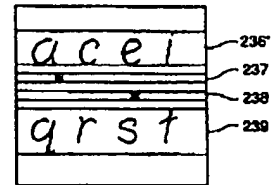
【図26 D】



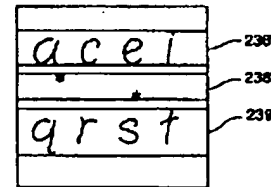
【図4A】



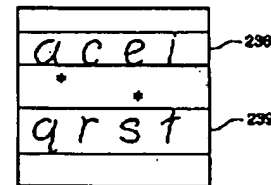
【図18B】



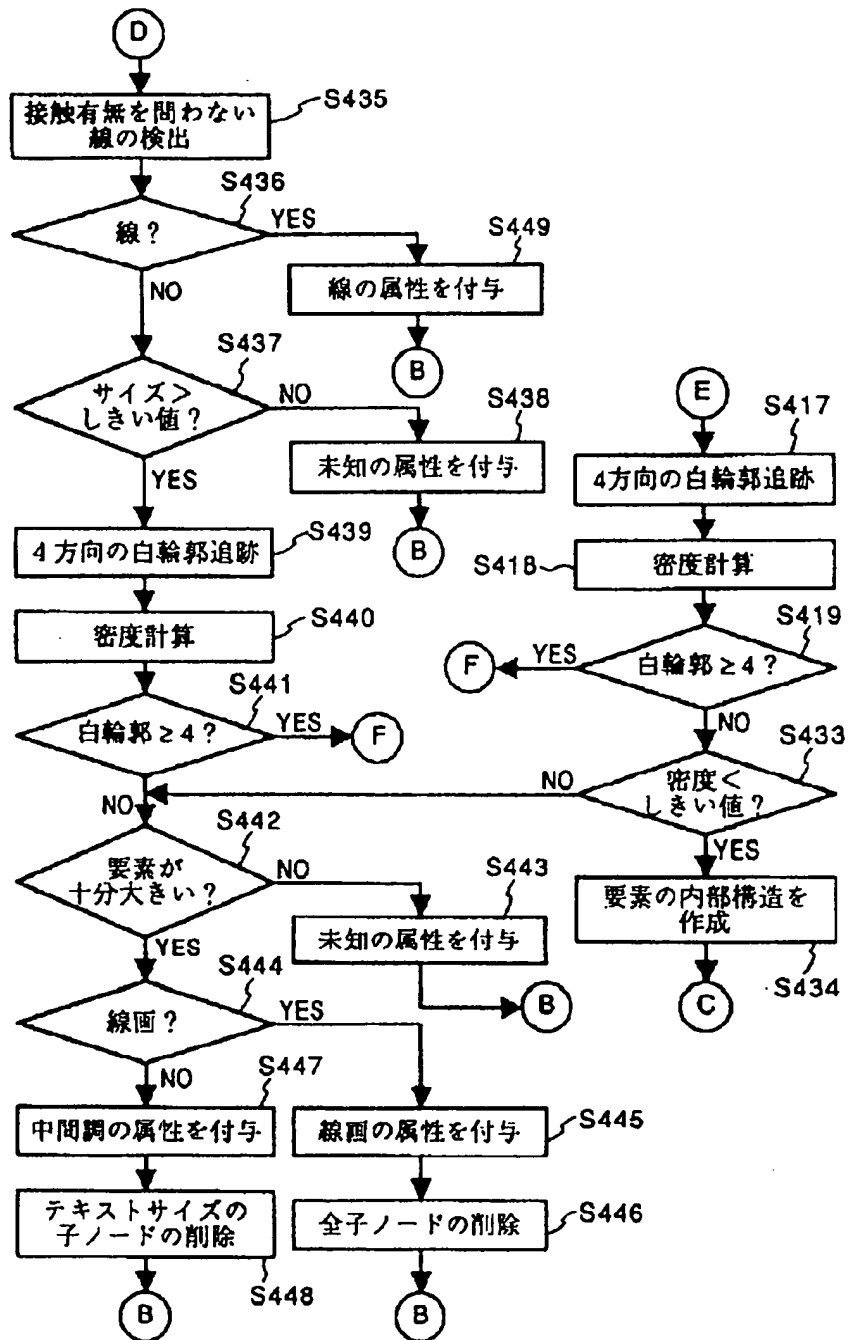
【図18C】



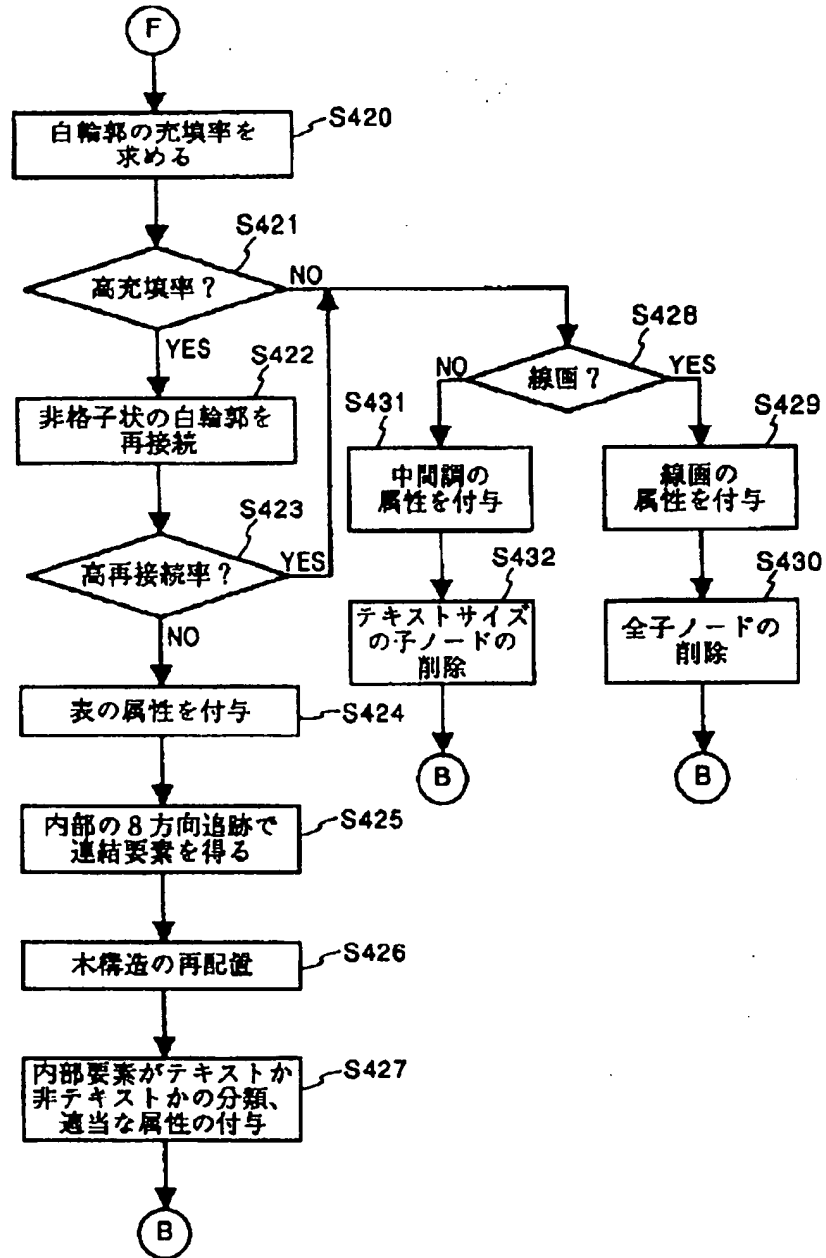
【図18D】



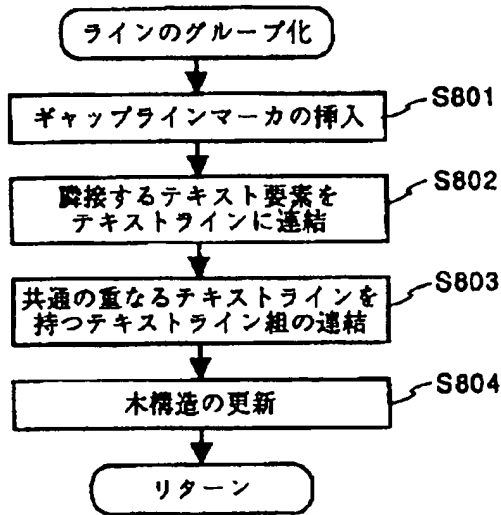
【図4B】



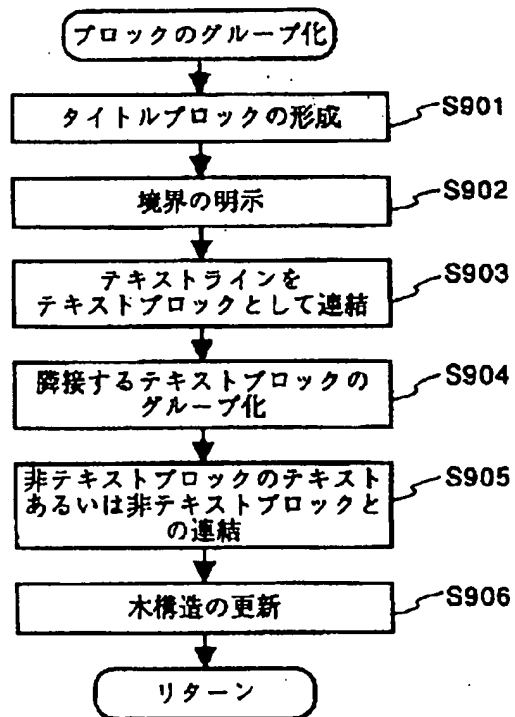
【図4C】



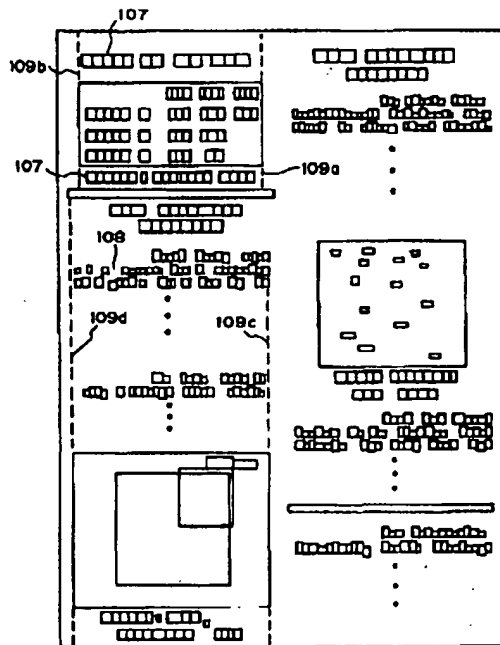
【図8】



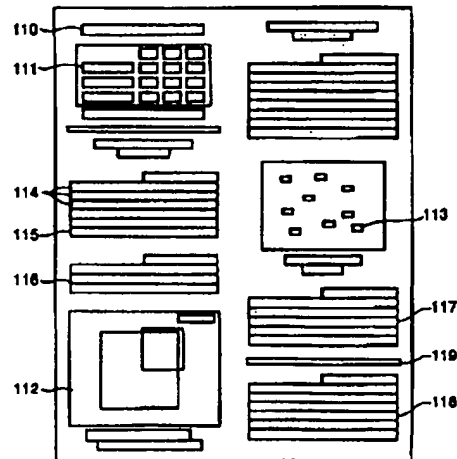
【図9】



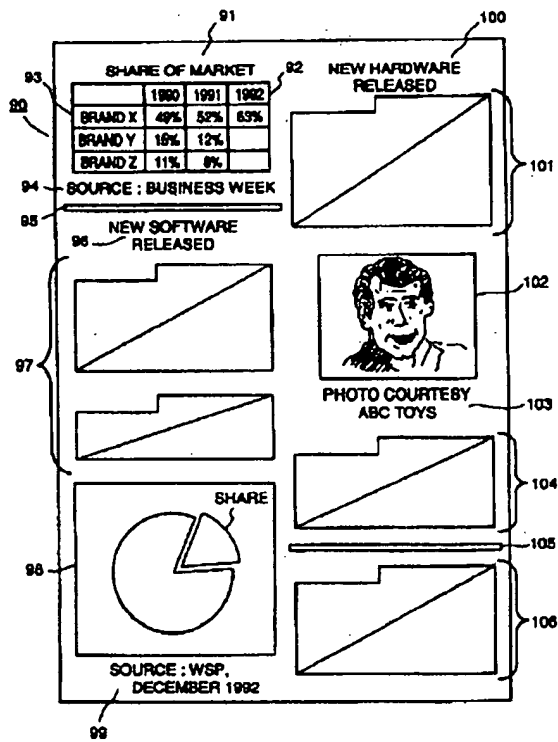
【図11】



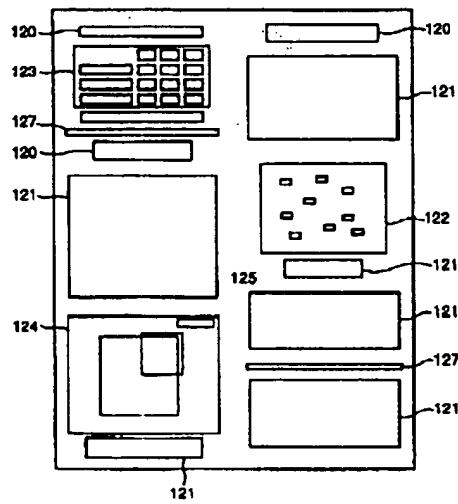
【図12】



【図10】



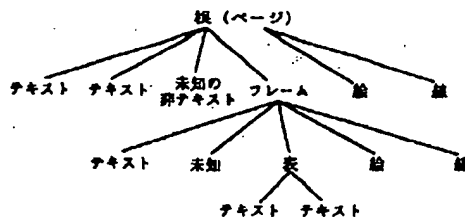
【図13】



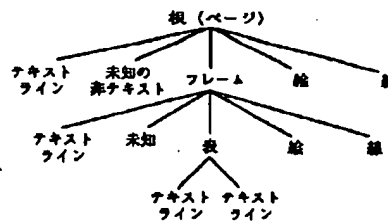
【図26A】

Satisfy

【図14】

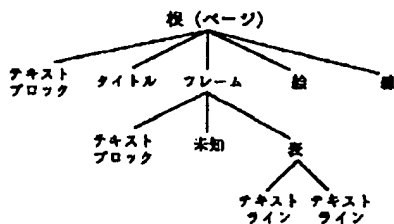


【図15】

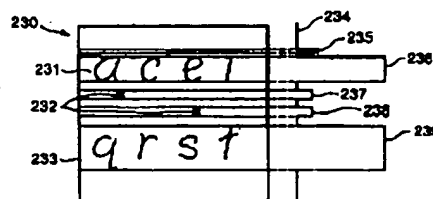


【図35】
Source Code For
Block Selection

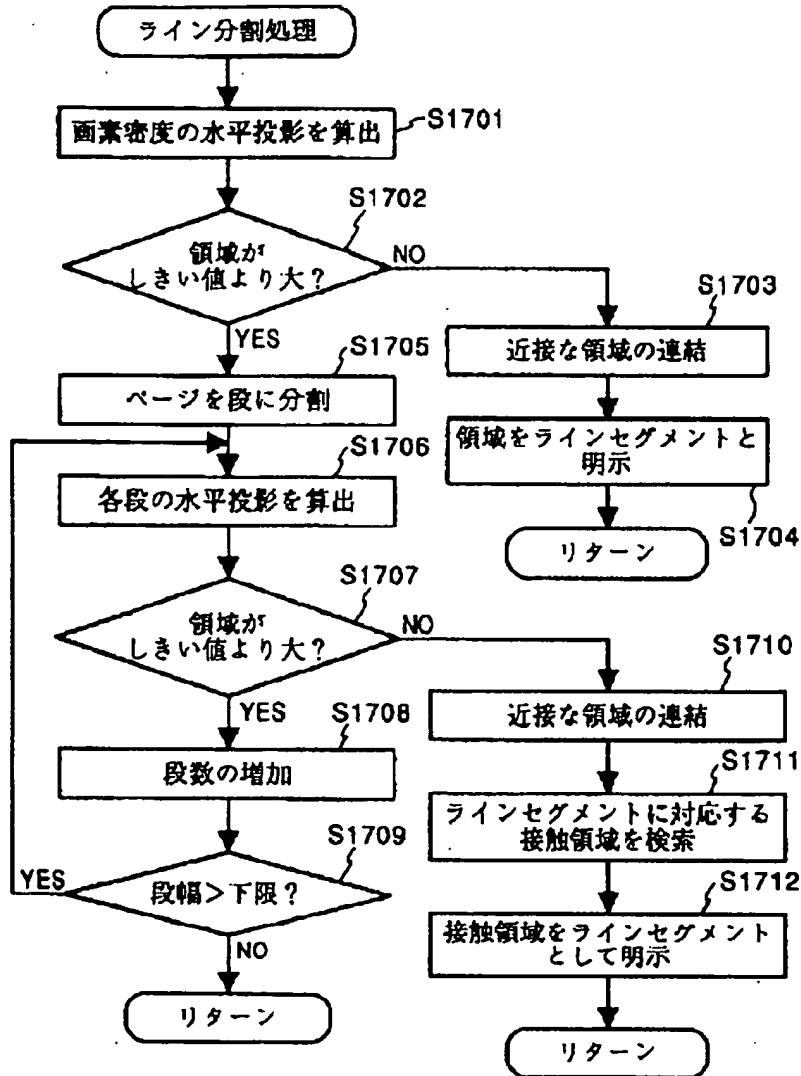
【図16】



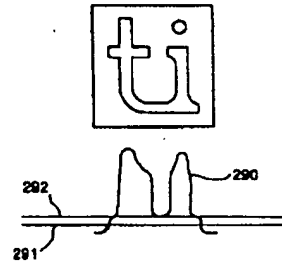
【図18A】



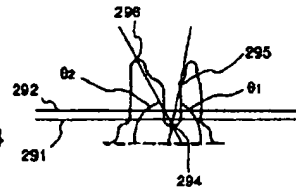
【図17】



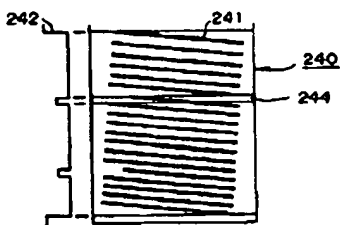
【図29A】



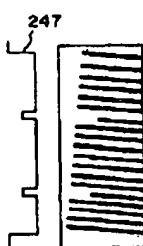
【図29B】



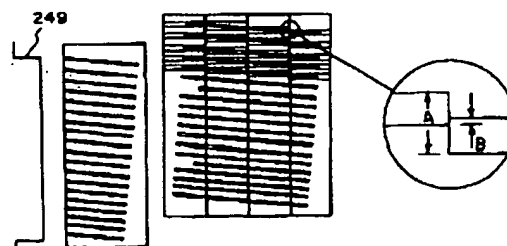
【図19A】



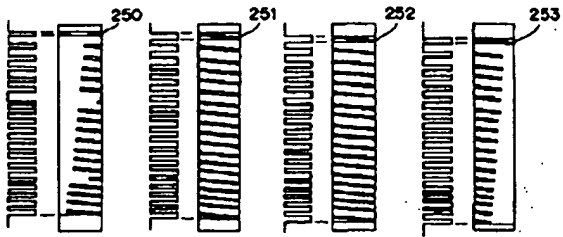
【図19B】



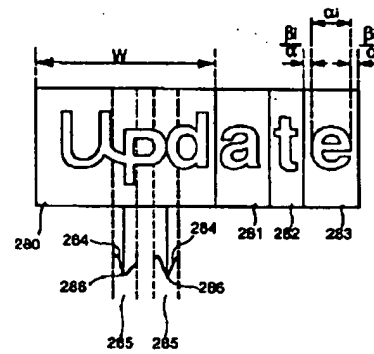
【図19D】



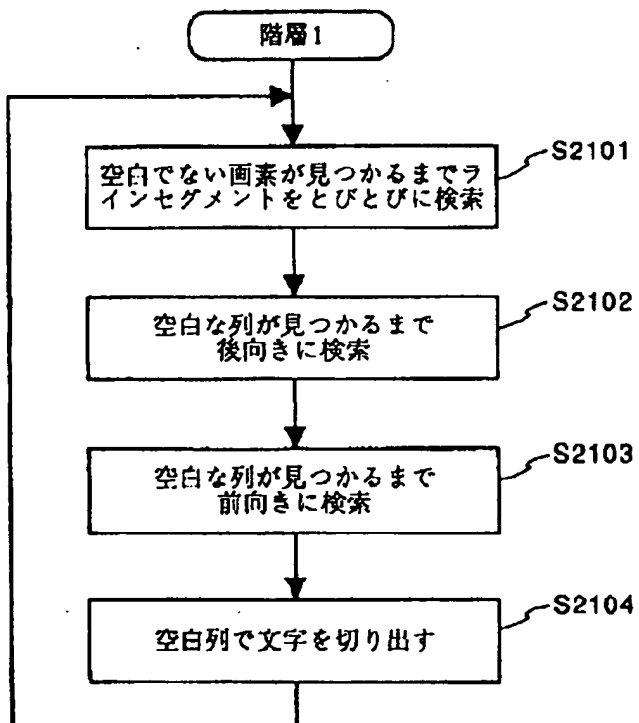
【図19C】



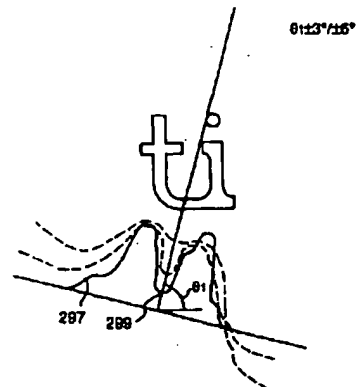
【図24】



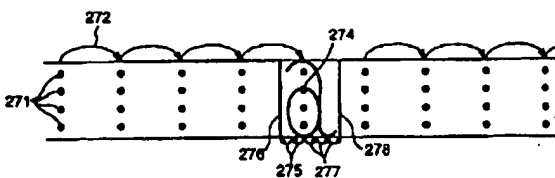
【図21】



【図29C】



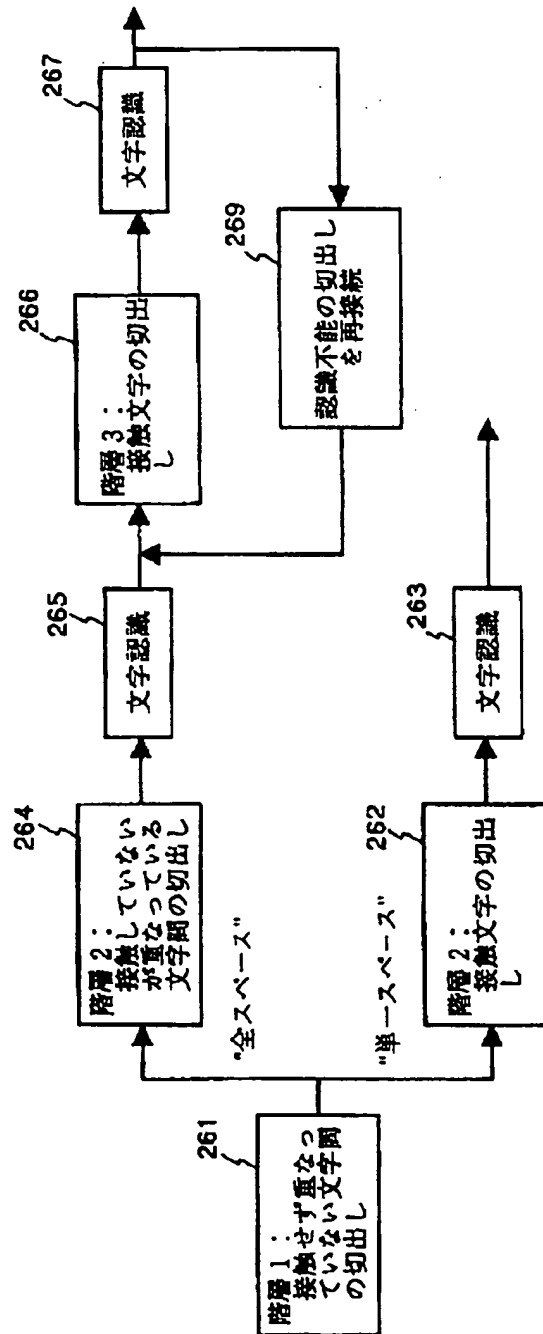
【図22】



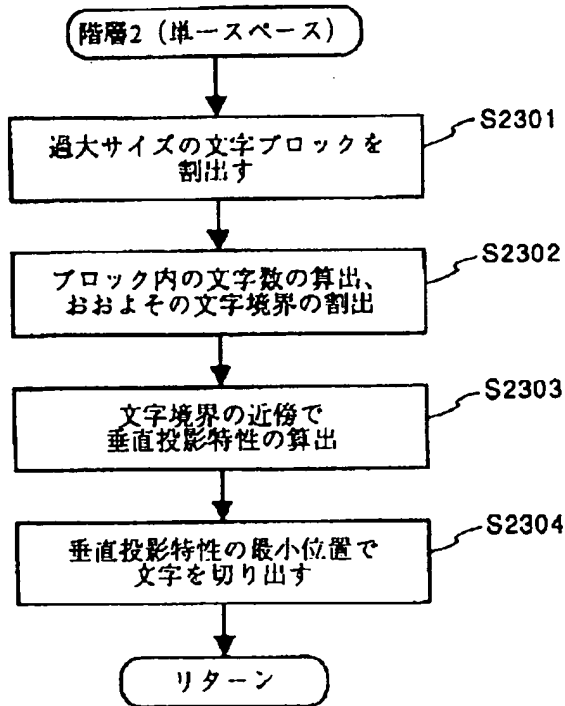
【図29D】



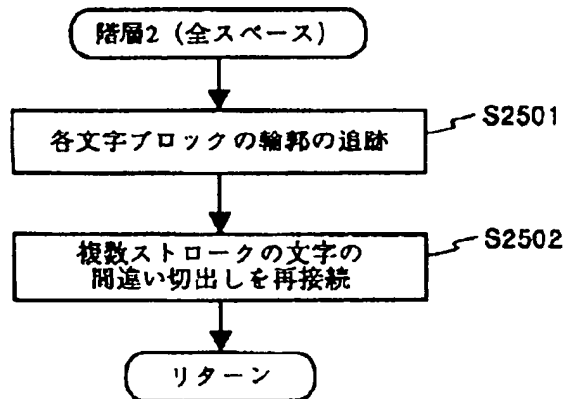
【図20】



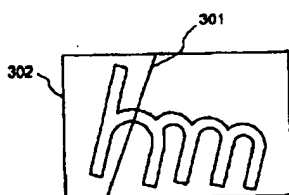
【図23】



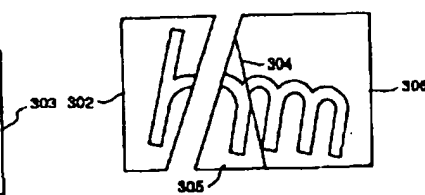
【図25】



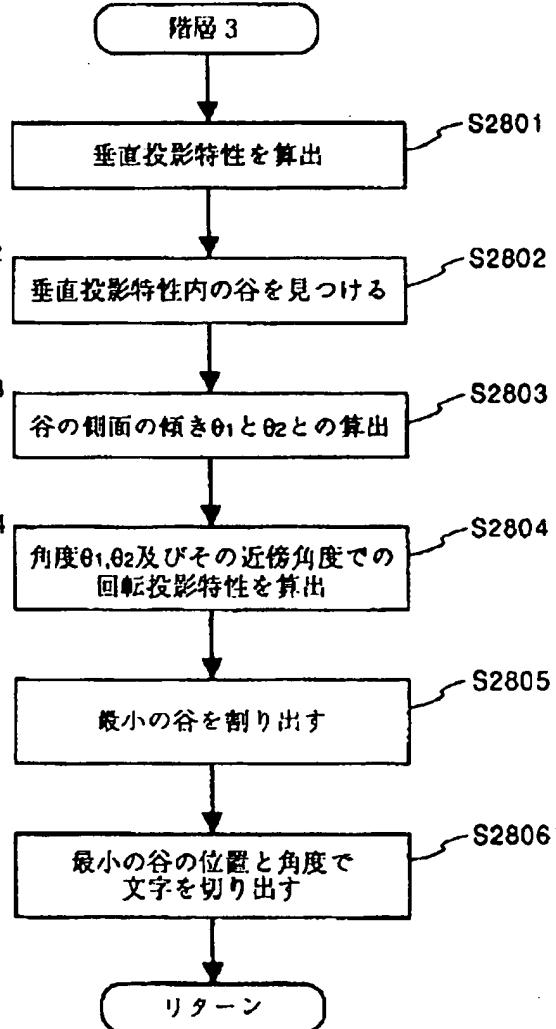
【図31A】



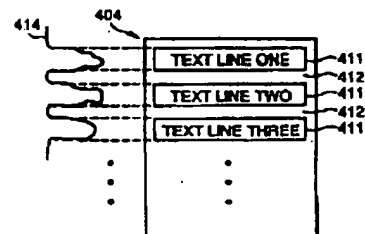
【図31B】



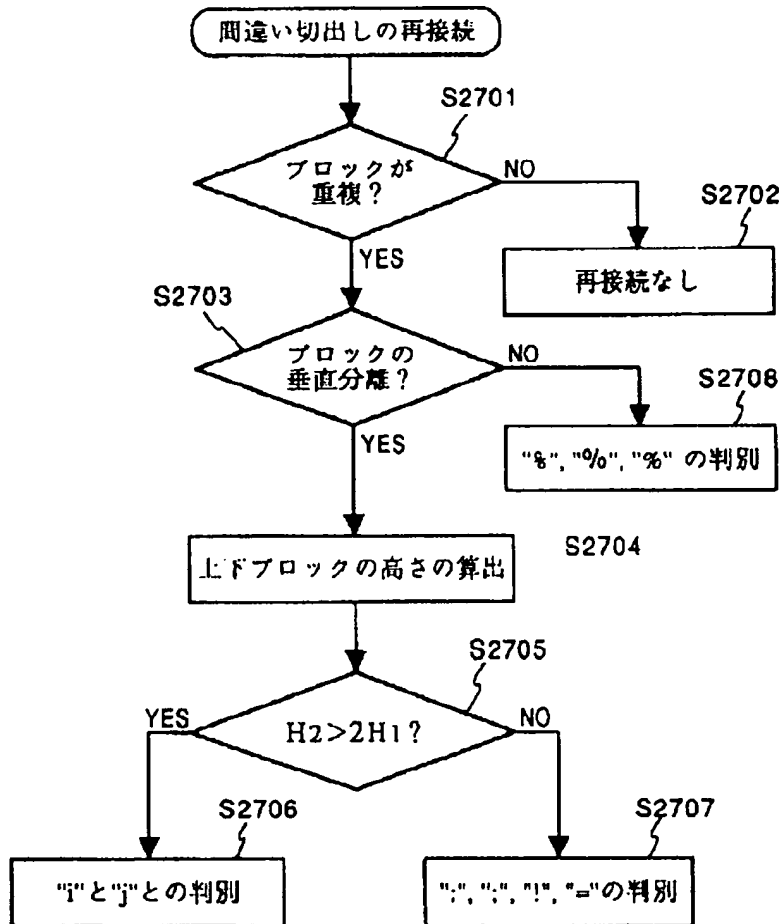
【図28】



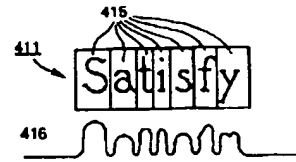
【図33A】



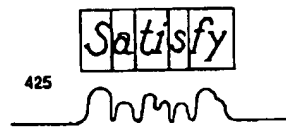
【図27】



【図34A】



【図34B】

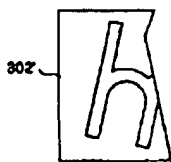


【図134】

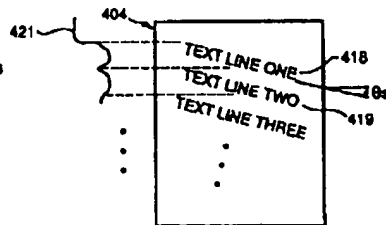
```

extern float *vector();
extern unsigned char *vector();
extern int *ivector();
extern double *dvector();
extern unsigned char *matrix();
extern float *matrix();
extern int **imatrix();
extern double **dmatrix();
extern void free_vector();
extern void free_ivector();
extern void free_dvector();
extern void free_matrix();
extern void free_imatrix();
extern void free_dmatrix();
  
```

【図31C】



【図33B】



【図129】

```

struct sectionblock *previous;
struct sectionblock *next;

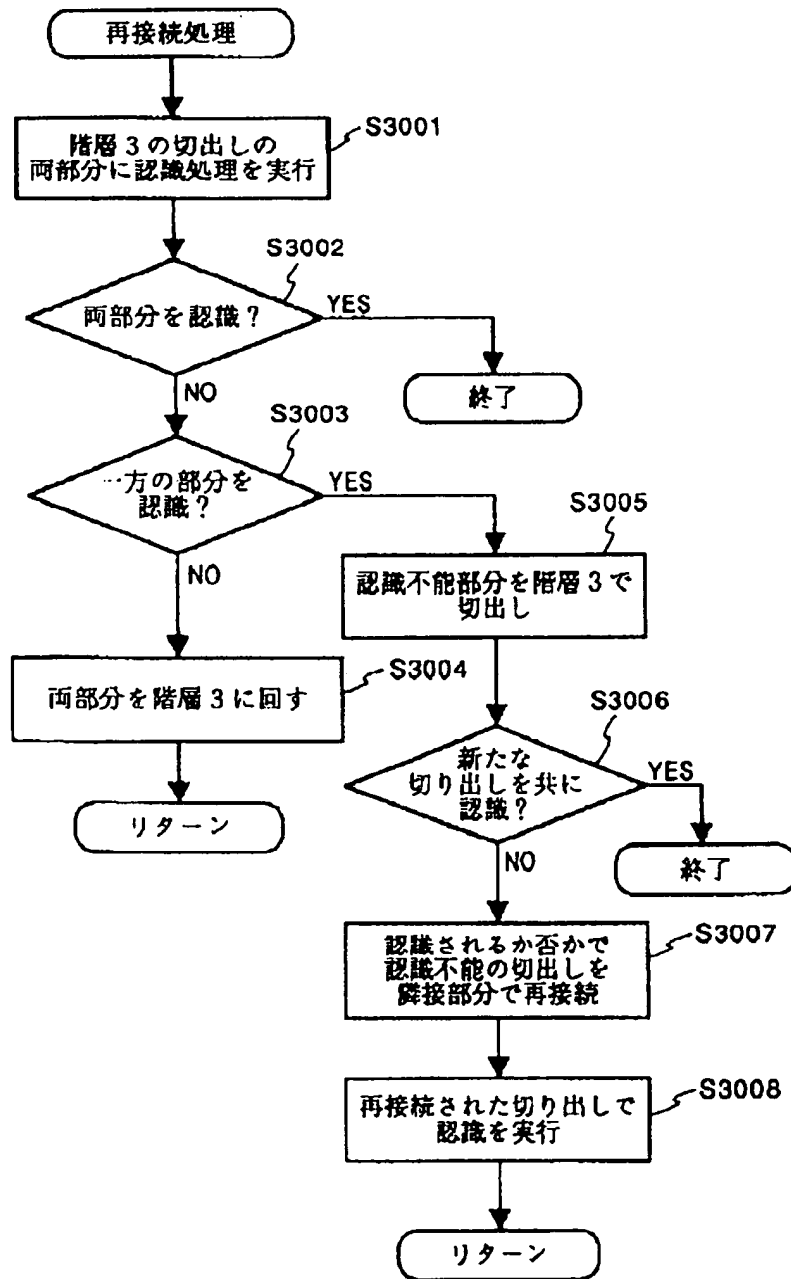
int upper_y, lower_y, left_x, right_x;
int width, length;
int lindex;
int outlineblock;
int ass;
int boundary_index;
  
```

【図163】

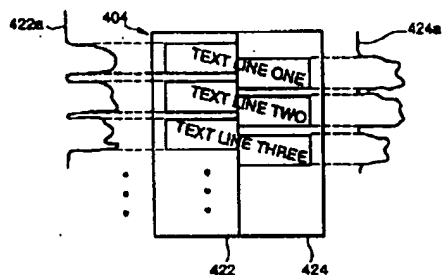
```

#define HISTOGRAM_THRESHOLD
charon = charon - MAX; // Get the max character
free_vector(histogram); // Deallocate the memory
  
```

【図30】



【図 3 3 C】



【图 6 2】

```
int width, height;
{
    int x, y, xdim, ydim, i;
    static fid = 33;

    init_graphics(wd);
    kd = kd_wm(wd);
    xdim = ((height-30)/(TOTALCOLOR));
    kd_getint(wd);
    for (i = 0; i < TOTALCOLOR; i++)
        kd_get_foreground(color_code[i], wd);
    kd_getfont(wd, fid);
    kd_string(wd, 30, (i+1)*xdim+20, color_label[i]);
}
kd_flush(wd);
```

【图 4-4】

[illegible]

【☒ 1 3 3】

```

#define R      0
#define G      1
#define B      2

int color_comp[3][3]={
    {0x43, 0xc3, 0x00},
    {0xc3, 0x00, 0x00},
    {0xc3, 0x00, 0xc3},
    {0xc3, 0xc3, 0xc3},
    {0xc3, 0xc3, 0xc3},
    {0xc3, 0xc3, 0xc3},
    {0xc3, 0xc3, 0xc3},
    {0xc3, 0xc3, 0xc3},
    {0xc3, 0xc3, 0xc3}
};

```

[illegible]

【图 4-6】

```
        _lhcCon->DisplayForm = DisplayForm;
        _lhcCon->DisplayFormClose_Func = DisplayFormClose_Func;
        _lhcCon->DisplayForm_DrawArea = DisplayForm_DrawArea;
    }

    static void _lhcFromContext (_lhcCon)
    {
        _lhcDisplayForm = *_lhcCon;

        DisplayForm = _lhcCon->DisplayForm;
        DisplayFormClose_Func = _lhcCon->DisplayFormClose_Func;
        DisplayForm_DrawArea = _lhcCon->DisplayForm_DrawArea;
    }

/* UNUSED */
static void activateCallback_DisplayFormClose_Func(HWND hwnd, LPARAM lParam, WPARAM wParam)
{
    widget_t *w;
    _lhcDisplayForm = *_lhcCon;
    List_Queue(&_lhcCon->List);

    w = widget_t;
    widget_t *w2;

    while(w != NULL)
    {
        w2 = widget_t;
        if(w2 == w)
            continue;
        w2->ActivateCallback(DisplayFormClose_Func, hwnd, lParam, wParam);
        w2->DeactivateCallback();
        w2->Destroy();
    }
}

void _lhcToContext(_lhcCon, Context)
```

【例 37】

[illegible]

【图 3 8】

[illegible]

【图 39】

[illegible]

【例 40】

[illegible]

【図42】

```

        _UnCon->FileSelectionDialog= FileSelectionDialog;
        _UnCon->FileSelectionBox= FileSelectionBox;
        _UnCon->parent= parent;
    }

    static void _UnFromContext(_UnCon)
    {
        _UnFileSelectionDialog = _UnCon;
        FileSelectionDialog= _UnCon->FileSelectionDialog;
        FileSelectionBox= _UnCon->FileSelectionBox;
        parent= _UnCon->parent;
    }

    /* ARGSUSED */
    static void cancelCallback_FileSelectionBox(UxWidget, UxContext, UxCallbackArg)
    {
        Widget      _UnWidget;
        _UnFileSelectionDialog *UnContext;
        int          _UnCallbackArg;
        {
            _UnWidget      UxThisWidget;
            UxThisWidget= UxWidgetToWidget(UxWidget);
            _UnFromContext(UxContext);

            {
                UxPopdownInterface(FileSelectionDialog);
            }
            _UnToContext(UxContext);
        }
    }

    /* ARGSUSED */
    static void okCallback_FileSelectionBox(UxWidget, UxContext, UxCallbackArg)
    {
        Widget      _UnWidget;
        _UnFileSelectionDialog *UnContext;
        int          _UnCallbackArg;
        {
            _UnWidget      UxThisWidget;
            UxThisWidget= UxWidgetToWidget(UxWidget);
            _UnFromContext(UxContext);

            {
                handle_tiff_handle;
                UxPopdownInterface(FileSelectionDialog);
                strcpy(filename, get_input_filename());
                printf("Open (%s)\n", filename);
                tiff_handle = UxCreateSubproc("/usr/local/bin/X11/imageview", filename,
                UnxpendFile);
                if (tiff_handle == -1) {
                    error_create_subproc();
                    return;
                }
                if (UxRunSubproc(tiff_handle, NULL) == -1) {
                    error_run_subproc();
                    return;
                }
                open_tiff_file(filename);
            }
            _UnToContext(UxContext);
        }
    }

```

【図48】

```

    }

    static void _UnFromContext(_UnCon)
    {
        _UnInfoForm = _UnCon;
        {
            InfoForm= _UnCon->InfoForm;
            InfoFormClose_FBO= _UnCon->InfoFormClose_FBO;
            InfoFormArea= _UnCon->InfoFormArea;
        }
    }

    /* ARGSUSED */
    static void act(va; callback_InfoFormClose_FBO(UxWidget, UxContext, UxCallbackArg)
    {
        Widget      _UnWidget;
        _UnInfoForm *UnContext;
        int          _UnCallbackArg;
        {
            _UnWidget      UxThisWidget;
            UxThisWidget= UxWidgetToWidget(UxWidget);
            _UnFromContext(UxContext);

            {
                UxPopdownInterface(InfoForm);
            }
            _UnToContext(UxContext);
        }
    }

```


—59—

[illegible]

【图 47】

[illegible]

【图 50】

[illegible]

—64—

[illegible]

【図53】

```

        printf("\nError in opening %s\n", filename);
        exit(1);
    }
    tagfile_position = 0;

    /* Read in header and make sure it's there! */
    i = getbytes(tagfile, tagfile_position);
    /* printf("i = %d\n", i); */
    if (i != 0x4949)
    {
        printf("... Error: TIFF file not in Intel format!\n");
        exit(1);
    }

    /* Read in Magic number
       getbytes(tagfile, tagfile_position);
       printf("i = %d\n", i); */
    if (i != 0x002a)
    {
        printf("... BAD Magic Number for TIFF file!\n");
        exit(1);
    }

    /* Find first file directory
       j = getbytes(tagfile, tagfile_position);
       validate_tagfile_position = j;
       {
           i = getbytes(tagfile, tagfile_position);
           break;
       }
       /* Tag in Directory
       i = getbytes(tagfile, tagfile_position);
       while (i != 0)
       {
           tag = getbytes(tagfile, tagfile_position);
           type = getbytes(tagfile, tagfile_position);
           value = getbytes(tagfile, tagfile_position);
           if ((i & 1) == 1) /* (tag & 32768) == 0 */
           {
               printf("Error: Tag length != 4!\n");
               exit(1);
           }
           printf("tag = %d\n", tag);
           switch (tag)
           {
               case 0x0000: /* MetadataType */
                   if (value != 0)
                   {
                       printf("... can't deal with non-standard MetadataType\n");
                       exit(1);
                   }
                   break;
               case 0x0001: /* MetadataType */
                   if (value != 1)
                   {
                       printf("... MetadataType not full resolution!\n");
                       exit(1);
                   }
                   break;
               case 0x0002: /* Width */
                   width = value;
                   break;
               case 0x0003: /* Length */
                   length = value;
                   break;
           }
       }

```

```

        case 0x0004: /* Bits per sample */
            if (value != 1)
            {
                printf("... Error: Bits/Sample must be 1!\n");
                exit(1);
            }
            break;
        case 0x0005: /* Compression */
            if (value != 1)
            {
                printf("Compression Method Error!\n");
                exit(1);
            }
            break;
        case 0x0006: /* Black is */
            break;
        case 0x0007: /* Thresholding tag */
            if (value != 1)
            {
                printf("Thresholding Tag = illegal value!\n");
                exit(1);
            }
            break;
        case 0x0011: /* Offset to raster data */
            raster_offset = value;
            break;
        case 0x0017: /* Type = value */
            break;
        case 0x0018: /* Break */
            break;
        case 0x0019: /* Break */
            break;
        case 0x001a: /* Break */
            break;
        case 0x001b: /* Break */
            break;
        case 0x001c: /* Break */
            break;
        case 0x001d: /* Break */
            break;
        case 0x001e: /* Break */
            break;
        case 0x001f: /* Break */
            break;
        case 0x0020: /* Break */
            break;
        case 0x0021: /* Break */
            break;
        case 0x0022: /* Break */
            break;
        case 0x0023: /* Break */
            break;
        case 0x0024: /* Break */
            break;
        case 0x0025: /* Break */
            break;
        case 0x0026: /* Break */
            break;
        case 0x0027: /* Break */
            break;
        case 0x0028: /* Break */
            break;
        case 0x0029: /* Break */
            break;
        case 0x002a: /* Break */
            break;
        case 0x002b: /* Break */
            break;
        case 0x002c: /* Break */
            break;
        case 0x002d: /* Break */
            break;
        case 0x002e: /* Break */
            break;
        case 0x002f: /* Break */
            break;
        case 0x0030: /* Break */
            break;
        case 0x0031: /* Break */
            break;
        case 0x0032: /* Break */
            break;
        case 0x0033: /* Break */
            break;
        case 0x0034: /* Break */
            break;
        case 0x0035: /* Break */
            break;
        case 0x0036: /* Break */
            break;
        case 0x0037: /* Break */
            break;
        case 0x0038: /* Break */
            break;
        case 0x0039: /* Break */
            break;
        case 0x003a: /* Break */
            break;
        case 0x003b: /* Break */
            break;
        case 0x003c: /* Break */
            break;
        case 0x003d: /* Break */
            break;
        case 0x003e: /* Break */
            break;
        case 0x003f: /* Break */
            break;
        case 0x0040: /* Break */
            break;
        case 0x0041: /* Break */
            break;
        case 0x0042: /* Break */
            break;
        case 0x0043: /* Break */
            break;
        case 0x0044: /* Break */
            break;
        case 0x0045: /* Break */
            break;
        case 0x0046: /* Break */
            break;
        case 0x0047: /* Break */
            break;
        case 0x0048: /* Break */
            break;
        case 0x0049: /* Break */
            break;
        case 0x004a: /* Break */
            break;
        case 0x004b: /* Break */
            break;
        case 0x004c: /* Break */
            break;
        case 0x004d: /* Break */
            break;
        case 0x004e: /* Break */
            break;
        case 0x004f: /* Break */
            break;
        case 0x0050: /* Break */
            break;
        case 0x0051: /* Break */
            break;
        case 0x0052: /* Break */
            break;
        case 0x0053: /* Break */
            break;
        case 0x0054: /* Break */
            break;
        case 0x0055: /* Break */
            break;
        case 0x0056: /* Break */
            break;
        case 0x0057: /* Break */
            break;
        case 0x0058: /* Break */
            break;
        case 0x0059: /* Break */
            break;
        case 0x005a: /* Break */
            break;
        case 0x005b: /* Break */
            break;
        case 0x005c: /* Break */
            break;
        case 0x005d: /* Break */
            break;
        case 0x005e: /* Break */
            break;
        case 0x005f: /* Break */
            break;
        case 0x0060: /* Break */
            break;
        case 0x0061: /* Break */
            break;
        case 0x0062: /* Break */
            break;
        case 0x0063: /* Break */
            break;
        case 0x0064: /* Break */
            break;
        case 0x0065: /* Break */
            break;
        case 0x0066: /* Break */
            break;
        case 0x0067: /* Break */
            break;
        case 0x0068: /* Break */
            break;
        case 0x0069: /* Break */
            break;
        case 0x006a: /* Break */
            break;
        case 0x006b: /* Break */
            break;
        case 0x006c: /* Break */
            break;
        case 0x006d: /* Break */
            break;
        case 0x006e: /* Break */
            break;
        case 0x006f: /* Break */
            break;
        case 0x0070: /* Break */
            break;
        case 0x0071: /* Break */
            break;
        case 0x0072: /* Break */
            break;
        case 0x0073: /* Break */
            break;
        case 0x0074: /* Break */
            break;
        case 0x0075: /* Break */
            break;
        case 0x0076: /* Break */
            break;
        case 0x0077: /* Break */
            break;
        case 0x0078: /* Break */
            break;
        case 0x0079: /* Break */
            break;
        case 0x007a: /* Break */
            break;
        case 0x007b: /* Break */
            break;
        case 0x007c: /* Break */
            break;
        case 0x007d: /* Break */
            break;
        case 0x007e: /* Break */
            break;
        case 0x007f: /* Break */
            break;
        case 0x0080: /* Break */
            break;
        case 0x0081: /* Break */
            break;
        case 0x0082: /* Break */
            break;
        case 0x0083: /* Break */
            break;
        case 0x0084: /* Break */
            break;
        case 0x0085: /* Break */
            break;
        case 0x0086: /* Break */
            break;
        case 0x0087: /* Break */
            break;
        case 0x0088: /* Break */
            break;
        case 0x0089: /* Break */
            break;
        case 0x008a: /* Break */
            break;
        case 0x008b: /* Break */
            break;
        case 0x008c: /* Break */
            break;
        case 0x008d: /* Break */
            break;
        case 0x008e: /* Break */
            break;
        case 0x008f: /* Break */
            break;
        case 0x0090: /* Break */
            break;
        case 0x0091: /* Break */
            break;
        case 0x0092: /* Break */
            break;
        case 0x0093: /* Break */
            break;
        case 0x0094: /* Break */
            break;
        case 0x0095: /* Break */
            break;
        case 0x0096: /* Break */
            break;
        case 0x0097: /* Break */
            break;
        case 0x0098: /* Break */
            break;
        case 0x0099: /* Break */
            break;
        case 0x009a: /* Break */
            break;
        case 0x009b: /* Break */
            break;
        case 0x009c: /* Break */
            break;
        case 0x009d: /* Break */
            break;
        case 0x009e: /* Break */
            break;
        case 0x009f: /* Break */
            break;
        case 0x00a0: /* Break */
            break;
        case 0x00a1: /* Break */
            break;
        case 0x00a2: /* Break */
            break;
        case 0x00a3: /* Break */
            break;
        case 0x00a4: /* Break */
            break;
        case 0x00a5: /* Break */
            break;
        case 0x00a6: /* Break */
            break;
        case 0x00a7: /* Break */
            break;
        case 0x00a8: /* Break */
            break;
        case 0x00a9: /* Break */
            break;
        case 0x00aa: /* Break */
            break;
        case 0x00ab: /* Break */
            break;
        case 0x00ac: /* Break */
            break;
        case 0x00ad: /* Break */
            break;
        case 0x00ae: /* Break */
            break;
        case 0x00af: /* Break */
            break;
        case 0x00b0: /* Break */
            break;
        case 0x00b1: /* Break */
            break;
        case 0x00b2: /* Break */
            break;
        case 0x00b3: /* Break */
            break;
        case 0x00b4: /* Break */
            break;
        case 0x00b5: /* Break */
            break;
        case 0x00b6: /* Break */
            break;
        case 0x00b7: /* Break */
            break;
        case 0x00b8: /* Break */
            break;
        case 0x00b9: /* Break */
            break;
        case 0x00ba: /* Break */
            break;
        case 0x00bb: /* Break */
            break;
        case 0x00bc: /* Break */
            break;
        case 0x00bd: /* Break */
            break;
        case 0x00be: /* Break */
            break;
        case 0x00bf: /* Break */
            break;
        case 0x00c0: /* Break */
            break;
        case 0x00c1: /* Break */
            break;
        case 0x00c2: /* Break */
            break;
        case 0x00c3: /* Break */
            break;
        case 0x00c4: /* Break */
            break;
        case 0x00c5: /* Break */
            break;
        case 0x00c6: /* Break */
            break;
        case 0x00c7: /* Break */
            break;
        case 0x00c8: /* Break */
            break;
        case 0x00c9: /* Break */
            break;
        case 0x00ca: /* Break */
            break;
        case 0x00cb: /* Break */
            break;
        case 0x00cc: /* Break */
            break;
        case 0x00cd: /* Break */
            break;
        case 0x00ce: /* Break */
            break;
        case 0x00cf: /* Break */
            break;
        case 0x00d0: /* Break */
            break;
        case 0x00d1: /* Break */
            break;
        case 0x00d2: /* Break */
            break;
        case 0x00d3: /* Break */
            break;
        case 0x00d4: /* Break */
            break;
        case 0x00d5: /* Break */
            break;
        case 0x00d6: /* Break */
            break;
        case 0x00d7: /* Break */
            break;
        case 0x00d8: /* Break */
            break;
        case 0x00d9: /* Break */
            break;
        case 0x00da: /* Break */
            break;
        case 0x00db: /* Break */
            break;
        case 0x00dc: /* Break */
            break;
        case 0x00dd: /* Break */
            break;
        case 0x00de: /* Break */
            break;
        case 0x00df: /* Break */
            break;
        case 0x00e0: /* Break */
            break;
        case 0x00e1: /* Break */
            break;
        case 0x00e2: /* Break */
            break;
        case 0x00e3: /* Break */
            break;
        case 0x00e4: /* Break */
            break;
        case 0x00e5: /* Break */
            break;
        case 0x00e6: /* Break */
            break;
        case 0x00e7: /* Break */
            break;
        case 0x00e8: /* Break */
            break;
        case 0x00e9: /* Break */
            break;
        case 0x00ea: /* Break */
            break;
        case 0x00eb: /* Break */
            break;
        case 0x00ec: /* Break */
            break;
        case 0x00ed: /* Break */
            break;
        case 0x00ee: /* Break */
            break;
        case 0x00ef: /* Break */
            break;
        case 0x00f0: /* Break */
            break;
        case 0x00f1: /* Break */
            break;
        case 0x00f2: /* Break */
            break;
        case 0x00f3: /* Break */
            break;
        case 0x00f4: /* Break */
            break;
        case 0x00f5: /* Break */
            break;
        case 0x00f6: /* Break */
            break;
        case 0x00f7: /* Break */
            break;
        case 0x00f8: /* Break */
            break;
        case 0x00f9: /* Break */
            break;
        case 0x00fa: /* Break */
            break;
        case 0x00fb: /* Break */
            break;
        case 0x00fc: /* Break */
            break;
        case 0x00fd: /* Break */
            break;
        case 0x00fe: /* Break */
            break;
        case 0x00ff: /* Break */
            break;
    }
}

```

【図54】

【図217】

Source Code for

Mono-Spaced ("Courier")
Segmentation

```

if (v == EOF)
{
    printf("\n Error: Premature End of File found in file (line%u)",
        mit(i));
}
if ((v & 128) == black*128)
    page->pixel(i)[j] = 0;
else
    page->pixel(i)[j] = 1;
v = v << 1;
}
}
fclose(tagfile);
return page;
}

```

【図86】

```

        release_block(tmp1);
        tmp1 = tmp11;
    }
    for (i = tmp1; i != NULL; i++)
    {
        tmp12 = tmp1->next;
        release_block(tmp1);
        tmp1 = tmp12;
    }
    free_rectblock(initial_block);
    return;
}

/* ..... Draw text components for display ..... */
/* ..... */

void block_construct(p, width, height, p_ori)
/* signed char **p;
   int width, height;
   unsigned char **p_ori;
   {
    int i, j;
    /* open file for storing block data */
    original_p = p_ori;
    reduce_factor = SCALE_RATIO;

    for (i = 0; i <= 8; i++)
        for (j = 0; j <= 8; j++)
            direct_status[i][j] = 0;

    block_file = fopen("block_file.dat", "w");

    text_length = 150.0/SCALE_RATIO;
    total_text = 0;

    /* the root block stands for the whole page image unit */
    root_rectblock = new_block(0,0);
    root_rectblock->width = width;
    root_rectblock->length = height;
    root_rectblock->upper_y = 0;
    root_rectblock->lower_y = height-1;
    root_rectblock->left_x = 0;
    root_rectblock->right_x = width-1;

    /* find blocks inside the page range */
    block_in_range(p, 0, 0, width-1, height-1, root_rectblock);
    printf("get out of block_in_range\n");
    rearrange_rectblock(root_rectblock);
    printf("get out of rearrange_rectblock\n");

    text_discriminate(root_rectblock);

    printf("get out of text_disc\n");

    firstlevelmontextsearch(p, root_rectblock);
    printf("get out of (firstlevelmontextsearch)\n");
    print_block(1, root_rectblock);

    (close(block_file);
    printf("before get out of block_construct\n");
}

```


【図56】

```

        height = (ymax-ymin+1);
    }
    print("width = %d height = %d\n", width, height);
    if (color == 0) {
        for (y = 0; y < height; y++)
            for (x = 0; x < width; x++) {
                xl = (int)((float)x/width*displayform);
                yl = (int)((float)y/height*displayform);
                if (p[y][x] != 0)
                    x_color[y][x] = 1;
            }
        print("after plot\n");
    } else {
        for (y = 0; y < height; y++)
            for (x = 0; x < width; x++) {
                xl = (int)((float)x/width*displayform);
                yl = (int)((float)y/height*displayform);
                if (p[y][x] != 0)
                    x_color[y][x] = 1;
            }
        print("after flush\n");
    }
    x_color[ymin, xmin, ymax, xmax];
}

// x_color[ymin, xmin, ymax, xmax];
// print("after flush\n");
// ..... display the rectblock .....
void display_rectblock(firstblock, wd)
struct block_rect *firstblock;
widget wd;
{
    struct block_rect *temp;
    int xl, yl, xl1, yl1;

    print("rectblock display of %d", x_color[ymin, xmin, ymax, xmax]);
    if ((firstblock == NULL) || (firstblock == NULL))
        return;
    else
        for (temp = firstblock; temp != NULL; temp = temp->next) {
            xl = (int)((float)temp->x/width*displayform);
            yl = (int)((float)temp->y/height*displayform);
            xl1 = (int)((float)temp->x1/width*displayform);
            yl1 = (int)((float)temp->y1/height*displayform);
            if (xl < xl1)
                xl = xl1;
            if (yl < yl1)
                yl = yl1;
            x_color[ymin, xmin, ymax, xmax];
        }
    for (temp = firstblock; temp != NULL; temp = temp->next) {
        xl = (int)((float)temp->x/width*displayform);
        yl = (int)((float)temp->y/height*displayform);
        xl1 = (int)((float)temp->x1/width*displayform);
        yl1 = (int)((float)temp->y1/height*displayform);
        if (xl < xl1)
            xl = xl1;
        if (yl < yl1)
            yl = yl1;
        x_color[ymin, xmin, ymax, xmax];
    }
}

```

```

        if (xl < xl1) xl = xl1;
        if (yl < yl1) yl = yl1;
        x_color[ymin, xmin, ymax, xmax];
    }
    if (temp->next != NULL)
        continue;
    else
        display_rectblock(temp, wd);
}

// ..... display the rectblock .....
// ..... display the rectblock .....
void display_rectblock(firstblock, wd)
widget wd;
{
    int xl, yl, xl1, yl1;

    x_color[ymin, xmin, ymax, xmax];
    for (temp = firstblock; temp != NULL; temp = temp->next) {
        xl = (int)((float)temp->x/width*displayform);
        yl = (int)((float)temp->y/height*displayform);
        xl1 = (int)((float)temp->x1/width*displayform);
        yl1 = (int)((float)temp->y1/height*displayform);
        if (xl < xl1)
            xl = xl1;
        if (yl < yl1)
            yl = yl1;
        x_color[ymin, xmin, ymax, xmax];
    }
    for (temp = firstblock; temp != NULL; temp = temp->next) {
        xl = (int)((float)temp->x/width*displayform);
        yl = (int)((float)temp->y/height*displayform);
        xl1 = (int)((float)temp->x1/width*displayform);
        yl1 = (int)((float)temp->y1/height*displayform);
        if (xl < xl1)
            xl = xl1;
        if (yl < yl1)
            yl = yl1;
        x_color[ymin, xmin, ymax, xmax];
    }
}

```



```

for (loop = firstblock->firstwhite; loop != NULL; )
{
    printf("white = %d\n", loop->numwhite);
    if (loop->numwhite == 0)
        loop = loop->next;
    else
        for (i = 0; i < loop->numwhite; i++)
            from_block[i] = *loop->white[i];
        loop->white[i] = 0;
        loop->numwhite--;
        loop->next = loop->next;
    }
}

free(struct whiteblock *loop->white);
}

free_white(loop);
}

printf("after free white, next = %d\n", loop);
loop = loop->next;
}

printf("after release white\n");
}

free_black(loop);

}

//===== free_section
//=====
void free_section(firstsection)
struct sectionblock *firstsection;
{
    struct sectionblock *loop, *temp;
    struct lineblock *loop2, *temp2;
    struct block_text *text;
    printf("got into free section\n");
    for (loop = firstsection->firstsubsection; loop != NULL; )
        temp = loop->next;
    printf("section = %d section = %d\n", loop, loop->indent);
    for (temp2 = loop->firstlineblock; temp2 != NULL; )
        loop2 = temp2->next;
    printf("loop = %d line = %d first = %d current = %d\n", loop,
           temp2->index, temp2->first_line_index, temp2->current_line_index);
    temp2->current_line_index = loop2->index;
    temp2->next_line_index = loop2->next;
    }

for (loop = loop->firstblock; loop != NULL; )
    temp = loop->next;
free_block(temp);
}

printf("after free rect, loop2 = %d loop = %d\n", loop2, loop);
loop2 = loop;
}

loop2 = loop;
loop2 = loop;
}

loop2->first_block = NULL;
printf("free line loop = %d loop2->index = %d undered temp2->next\n",
       loop, temp2->index, temp2->next_line_index, temp2->next_line_index);
}

free(struct lineblock *temp);
}

printf("after free line loop = %d\n", loop);
loop = loop->next;
}

```

[illegible]

—72—

[illegible]

【図64】

```

/* Allocate a double matrix with range [n1..n2] [n1..n2]. */
double **matrix(n1, n2, n1, n2) {
    int n1, n2, n1, n2;
    {
        int i;
        double **m;
        m = (double **) malloc((n2-n1+1)*sizeof(double));
        if (!m) perror("allocation failure 1 in matrix");
        m += n1;
        for (i = n1; i <= n2; i++) {
            m[i] = (double *) malloc((n2-n1+1)*sizeof(double));
            if (!m[i]) perror("allocation failure 2 in matrix");
            m[i] += n1;
        }
        return m;
    }
}

/* deallocate a float vector */
void free_vector(float *v, n1, n2) {
    int n1, n2;
    free((float *) (v-n1));
}

/* deallocate an int vector */
void free_vector(int *v, n1, n2) {
    int *v, n1, n2;
    free((int *) (v-n1));
}

/* deallocate a double vector */
void free_vector(double *v, n1, n2) {
    double *v, n1, n2;
    free((double *) (v-n1));
}

/* deallocate a float matrix */
void free_matrix(float m, n1, n2, n1, n2) {
    float **m, n1, n2, n1, n2;
    {
        int i;
        for (i = n1; i <= n2; i++) free((float *) (m[i]-n1));
        free((float *) (m-n1));
    }
}

/* deallocate an int matrix */
void free_matrix(int m, n1, n2, n1, n2) {
    int **m, n1, n2, n1, n2;
    {
        int i;
        for (i = n1; i <= n2; i++) free((int *) (m[i]-n1));
        free((int *) (m-n1));
    }
}

```

【例 6 5】

[illegible]

—78—

[illegible]

[図68]

```

void set_outline(x, y, nblock)
int
struct waitblock *nblock;
{
    if (x < nblock->right_x)
        nblock->right_x = x;
    else if (x < nblock->left_x)
        nblock->left_x = x;
    if (y > nblock->lower_y)
        nblock->lower_y = y;
    else if (y < nblock->upper_y)
        nblock->upper_y = y;
}

/*----- Set up the extreme point for the searched block -----*/
void set_waiting_outline(x, y, nblock)
int x, y;
struct block *nblock;
{
    if (x < nblock->right_x)
        nblock->right_x = x;
    else if (x < nblock->left_x)
        nblock->left_x = x;
    if (y > nblock->lower_y)
        nblock->lower_y = y;
    else if (y < nblock->upper_y)
        nblock->upper_y = y;
}

/*----- Allocate into memory -----*/
int nl, nh;
int **;
{
    if (nl < {int **} malloc((nh-nl+1)*sizeof(int **)) == NULL) {
        printf("allocation error. in inpointer");
        exit(1);
    }
    v = v+1;
    return v;
}

/*----- Transfer outline data into pair -----*/
void put_outline_into_pair(nblock, ylist)
struct block *nblock;
struct outline *ylist;
{
    int i, k;
    struct _outline *tmp;
    nblock->pairright = (int *)malloc(sizeof(nblock->pairright), nblock->lower_y);
    nblock->pairleft = (int *)malloc(sizeof(nblock->pairleft), nblock->upper_y);
    for (i = nblock->upper_y; i < nblock->lower_y; i++)

```


—82—

[illegible]

~~—85—~~

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

[illegible]

—88—

[illegible]

【图 7 8】

[illegible]

【図79】

```

return block->next; "newblock";
int i, y;
printf("get base block split x = %d y = %d cover = %d first = %d stream = %d\n",
    tap->split_x, tap->split_y, tap->cover, tap->first, tap->stream);
}

newblock = new_block(0,0);
newblock->split_x = tap->split_x;
newblock->split_y = tap->split_y;
newblock->upper_y = tap->upper_y;
newblock->lower_y = tap->lower_y;
newblock->right_x = tap->right_x;
newblock->left_x = tap->left_x;
for (y = newblock->upper_y; y < newblock->lower_y; y++)
    for (x = 1; x < tap->split_x[0]; x++)
        if (newblock->right_x < tap->split_x[0])
            newblock->right_x = tap->split_x[0];
for (y = newblock->lower_y; y < newblock->upper_y; y++)
    for (x = 1; x < tap->split_x[0]; x++)
        if (newblock->left_x < tap->split_x[0])
            newblock->left_x = tap->split_x[0];
newblock->length = newblock->lower_y - newblock->upper_y;
newblock->point = tap->point; newblock->length = tap->length;
newblock->y = gap2;

newblock->x = newblock->right_x;
for (i = 1; i < newblock->length; i++)
    if (tap->split_x[newblock->upper_y][i] < newblock->x)
        newblock->x = tap->split_x[newblock->upper_y][i];
newblock->in_group = tap->in_group;
newblock->next = tap->next;

tap->lower_y = gap1;
tap->length = tap->lower_y - tap->upper_y;
tap->point = tap->point; newblock->point;
add_junction(TEXT, newblock, head);

/* print("get out of block split in") */
/* print("get out of block split in") */

/* ..... split the line connected component if same gap between .....
   ..... split the line connected component if same gap between ..... */
int tap_split(newblock, gap, len)
struct block *tap, *head;
float len;
{
    int i, j, total, split;
    int upper, lower, left, right;
    int gap1, gap2;
    printf("get into tap split len = %f\n", len);
}

```


[図 8 2]

```

        struct block_rect *tmp, *tmpnext;
        float
            tmp1 = (head->avg_height == 0.0) ? cont_length / head->avg_height;
        for (tmp = first->firstsubblock; tmp->AVUL; ) {
            if (tmp->cover == COVERED)
                if (!float)tmp->length == tmp->TEXT_SCALE)
                    tmpnext = tmp->next; at(tmp->first->head->width/width_RATIO);
                    transfer_block(TEXT, TEXT, tmp, first, head);
                    tmp = tmpnext;
                }
            else {
                if (tmp->length > tmp->TEXT_SCALE)
                    if (float)tmp1 < (first, tmp, tmp1) continue;
                    tmpnext = tmp->next;
                    transfer_block(TEXT, NONTEXT, tmp, first, head);
                    tmp = tmpnext;
                }
            } else tmp = tmp->next;
        }
    }
}

/*
 * Recover the covered content block into original head block
 */
void main_recover(head, first)
struct block_rect *head, *first;
{
    struct block_rect *tmp, *tmpnext;
    float
        tmp1 = (head->avg_height == 0.0) ? cont_length / head->AVUL;
    for (tmp = first->firstsubblock; tmp->AVUL; ) {
        if (tmp->cover == COVERED)
            tmpnext = tmp->next;
            if (!float)tmp->length == tmp->TEXT_SCALE)
                tmpnext = tmp->next; at(tmp->first->head->width/width_RATIO);
                transfer_block(TEXT, TEXT, tmp, first, head);
                tmp = tmpnext;
            }
        else {
            if (tmp->length > tmp->TEXT_SCALE)
                if (float)tmp1 < (first, tmp, tmp1) continue;
                tmpnext = tmp->next;
                transfer_block(TEXT, NONTEXT, tmp, first, head);
                tmp = tmpnext;
            }
        } else tmp = tmp->next;
    }
}

void main_recover(head, first)
{
    /*
     * Recover the covered block into original head block
     */
    void main_recover(head, first);
    struct block_rect *head, *first;
    void main_recover(head, first);
}

```

[illegible]

[illegible]

【图 8-5】

[illegible]

【图 8-7】

[illegible]

—98—

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

【図 8 9】

```

void add_block_into_line(rect_block, head_block, line_block)
struct block *rect_block, *head_block,
struct line_block *line_block;
{
    struct block *top_block;
    remove_block(rect_block, head_block);
    head_block->next_block = rect_block;
    if (line_block->first_block == NULL) {
        line_block->first_block = rect_block->current_block = rect_block;
        rect_block->previous = rect_block->next = NULL;
        line_block->right = rect_block->right;
        line_block->left = rect_block->left;
        line_block->up = rect_block->up;
        line_block->down = rect_block->down;
    } else {
        line_block->first_block->previous = rect_block;
        rect_block->next = line_block->first_block;
        line_block->first_block->previous = rect_block;
        if (rect_block->right == rect_block->next) {
            line_block->right = rect_block->right;
        }
        if (rect_block->left == rect_block->left) {
            line_block->left = rect_block->left;
        }
        if (rect_block->up == rect_block->up) {
            line_block->up = rect_block->up;
        }
        if (rect_block->down == rect_block->down) {
            line_block->down = rect_block->down;
        }
        line_block->current_block = rect_block;
    }
}

/* test if a gap between two rect_block */
.....
int is_gap(rect_block, block, head_block)
struct block *rect_block, *block, *head_block;
{
    int count1, count2;
    int line, maxcount, mincount;

    if (block->right == rect_block->right) return 0;
    bound1 = min(block->up, block->down);
    bound2 = max(block->up, block->down);
    for (count1 = block->right; count1 < block->left; count1++) {
        if (rect_block->up == bound1) {
            print("up", rect_block->up);
            count1++;
        }
    }
    if (count1 == (bound2 - bound1)) return 1;
}

int = (MURPHY/SCALE_FACTOR + 1) * (count1 + 1) * MURPHY/SCALE_FACTOR;
mincount = 0;
maxcount = 0;
}

```


【图 9-1】

[illegible]

-102-

[illegible]

—103—

[illegible]

【例 9 4】

[illegible]

—105—

[illegible]

【図96】

```

                                tapline=tapline->next();
for ( i = 0; i <= (level-1)*8; i++)
    fprintf(line_file, " ");
fprintf(line_file, "td content index = %d ", tapline, tapline->index);
fprintf(line_file,
"xtd ytd lstd wtd att = %s block indexlstd indexlstd y = %d i = %d r = %d\n",
    tapline->left_x, tapline->upper_y, tapline->length,
    tapline->width, tapline->att, tapline->(first_block->index,
    tapline->current_block->index, tapline->group, tapline->l_fcaption,
    tapline->r_fcaption);
    print_ln(nextlevel, tapline);
}

/***** Open the file for the output of line message *****/
void print_lineblock(firstline)
struct lineblock *firstline;
{
    line_file = fopen("line_file.dat", "w");
    print_lnfl, firstline);
    fclose(line_file);
}

```

【図118】

```

                                tapsection = tapsection->next();
if ((tapsection->att&(TITLE|HACKID|VOLUME|AF|TRUE))) continue;
else
    ini_section_block(tapsection, tapsection->firstlineblock, p,
                                column);
}

/***** Print the line information into a file *****/
void print_sec(level, firstsec)
int level;
struct sectionblock *firstsec;
{
    int nextlevel, i;
    struct sectionblock *tapsubsec, *tapsec;
    nextlevel = level+1;
    for (tapsubsec = firstsec->firstsubsec; tapsubsec != NULL;
        tapsubsec = tapsubsec->next){
        for ( i = 0; i <= (level-1)*8; i++)
            printf(" ");
        printf("index = %d ", tapsubsec->index);
        printf("xtd ytd lstd wtd att = %s\n",
            tapsubsec->left_x, tapsubsec->upper_y,
            tapsubsec->length, tapsubsec->width, tapsubsec->att);
    }
    for (tapsec = firstsec->firstmainsection; tapsec != NULL;
        tapsec=tapsec->next){
        for ( i = 0; i <= (level-1)*8; i++)
            printf(" ");
        printf("content index = %d ", tapsec->index);
        printf("xtd ytd lstd wtd att = %s\n",
            tapsec->left_x, tapsec->upper_y, tapsec->length,
            tapsec->width, tapsec->att);
        print_sec(nextlevel, tapsec);
    }

/***** Open the file for the output of line message *****/
void print_secblock(firstsec)
struct sectionblock *firstsec;
{
    print_sec(1, firstsec);
}

```

【図137】

Source Code for Proportional Spaced Segmentation

—107—

```

for (i = 0; i < n; i++)
    v[i].column_no = 0;
return v;
}

//===== allocation section pointers
//=====

struct sectionblock *new_sectionpointer(int, int)
{
    int n1, n2;

    struct sectionblock **v;

    v = (struct sectionblock **) malloc((n1-n2+1)*sizeof(struct sectionblock));
    if (!v) {
        printf("error in the allocation of sectionpointer\n");
        exit(1);
    }
    v = v+n1;

    return v;
}

//===== sort line block pointer array
//=====

void sort_linepointer(int n, int)
{
    struct lineblock **line;
    int pos;

    int i, j, stop;
    struct lineblock *tmp;

    for (i = 0; i < n-1; i++) {
        for (j = i, stop = 0; j < n-1; j++)
            if (line[j]-segment_j > line[j+1]-segment_j) {
                tmp = line[j+1];
                line[j+1] = line[j];
                line[j] = tmp;
                stop = 1;
            }
        if (stop == 0) break;
    }
}

//===== line block sorting
//=====

struct lineblock *sort_linepointer(struct lineblock, struct lineblock)
{
    int *n1, n2;

    struct lineblock *tmp;
    struct lineblock **line;
    int i, j, stop;

    *n1 = 0;
    for (tmp = first_lineblock->first_lineblock; tmp != NULL; tmp = tmp->next)
        (*n1)++;
    for (tmp = first_lineblock->first_lineblock; tmp != NULL; tmp = tmp->next)
        (*n2)++;

    line = new_linepointer(0, (*n1)-1);
}

```

—108—

```

for (i = 0; temp = firstlineof-nextsubline; temp != NULL; temp = temp->next)
    link_x[i++] = temp;

/* exit setting */
if (firstlineof-nextsubline != NULL) {
    printf("error in setting\n");
}

for (temp = firstlineof-nextsubline; temp != NULL; temp = temp->next)
    link_x[i++] = temp;

sort_topofnextline(link_x, "newline");
return link_x;
}

/*..... section block setting */
struct sectionblock *next_section(first, num)
int num;
{
    struct sectionblock *top;
    struct sectionblock *first;
    int i, j, stop;

    num = 0;
    first = first->firstsubsection; temp = NULL; temp = temp->next;
    link_x[i++] = temp;

    for (j = 0; temp = first->firstsubsection; temp != NULL; temp = temp->next)
        link_x[i++] = temp;

    for (i = 0; i = (num-1); i++) {
        for (j = 0; temp = temp->firstsubsection; temp != NULL; temp = temp->next)
            link_x[i+j] = temp;
        temp = link_x[i+j];
        link_x[i+j-1] = link_x[j];
        link_x[i] = temp;
        temp = j;
    }

    if (temp != 0) break;
}

return link_x;
}

/*..... Allocate new line block */
struct sectionblock *new_sectionblock(int, int)
int n1, n2;
{
    struct sectionblock *v;
    v = (struct sectionblock *) malloc(sizeof(struct sectionblock));
    if (v) {
        printf("Allocation error: section block\n");
        printf(".....");
    }
}

```

[圖 9 9]

```

    if (current_line->first_lineblock == NULL)
    {
        current_line->first_lineblock = line;
        if (current_line->current_lineblock == line)
        {
            current_line->current_lineblock = NULL;
        }
        else
        {
            current_line->first_lineblock->next = line;
            if (current_line->current_lineblock == NULL)
            {
                current_line->current_lineblock = line;
            }
            else
            {
                line->previous = current_line->current_lineblock;
                current_line->current_lineblock->next = line;
            }
        }
        line->previous = NULL;
        line->next = NULL;
    }

    // Put lineblock into section
    .....
    void put_section(line, current_line)
    {
        struct lineblock *line;
        struct section *section;
        if (current_line->first_lineblock == NULL)
        {
            current_line->first_lineblock = line;
            current_line->current_lineblock = line;
            line->previous = NULL;
            line->next = NULL;
        }
        else
        {
            current_line->current_lineblock->next = line;
            line->previous = current_line->current_lineblock;
            current_line->current_lineblock = line;
            line->next = NULL;
        }
    }

    .....
    struct lineblock *line;
    struct section *section;
    v = (struct section *) malloc(sizeof(struct section));
    if (!v)
    {
        printf("allocation error in section\n");
        exit(1);
    }
    v->current = NULL;
    v->previous = NULL;
    v->next = NULL;
    v->first_lineblock = NULL;
    return v;
}

.....
void sort_section(struct)
{
    struct lineblock *line;
    struct section *section;
    .....
    void sort_section(struct)
    {
        struct lineblock *line;
        struct section *section;
        .....
}

```


-112-

[illegible]

-113-

[illegible]

【図106】

```

segment(current_seg).begin_column = vector(i, columnset);
segment(current_seg).end_column = vector(i, columnset);
for (i = 0; i < columnset; i++)
    segment(current_seg).begin_column(i) = gapend(i);
    segment(current_seg).end_column(i) = gapend(i);
}

for (j = 0; j < segment(current_seg).columnset; j++)
    printf("%d %d %d", segment(current_seg).begin_column(j),
           segment(current_seg).end_column(j));
    printf("\n");
}

for (tapline = section(columnset).firstlineblock, tapline != NULL,
    for (k = 0; k < columnset; k++)
        tapline = tapline->next;
        if (tapline->right > gapend(k))
            break;
        for (tap = section(k).firstlineblock, tap != NULL;
            for (i = 0; i < columnset; i++)
                if ((current_seg).begin_column(i) < tap->lower_y, tapline->upper_y,
                    (current_seg).end_column(i) > tap->upper_y, tapline->lower_y, tap->upper_y,
                    tap->lower_y > 0))
                    tapline = tap->next;
            if (tapline != NULL)
                if (tapline->right > gapend(k))
                    break;
            else
                put_section(tap, section(columnset));
            tap = tapline;
            continue;
        }
        tap = tap->next;
    }

for (k = 0; k < columnset; k++)
    for (i = 0; i < columnset; i++)
        if ((tapline->right_k < gapbegin(k+1)) &&
            (tapline->right_k > gapend(k)))
            break;
        if (k == columnset)
            for (tap = section(k).firstlineblock, tap != NULL;
                for (i = 0; i < columnset; i++)
                    if ((current_seg).begin_column(i) < tap->lower_y, tapline->upper_y,
                        (current_seg).end_column(i) > tap->upper_y, tapline->lower_y, tap->upper_y,
                        tap->lower_y > 0))
                        tapline = tap->next;
                if (tapline != NULL)
                    if (tapline->right_k < gapbegin(k+1)) &&
                        (tapline->right_k > gapend(k)))
                        break;
                else
                    put_section(tap, section(columnset));
                tap = tapline;
                continue;
            }
            tap = tap->next;
        }

for (i = 0; i < columnset; i++)
    printf("section %d %d", i);
}

```

-117-

-118-

[illegible]

【図110】

```

    (top->right_x == min(top->right_x, bottom->right_x))
    return 0;

    if ((!(overlap(top->upper_x, top->lower_x,
        min(top->upper_x, bottom->upper_x),
        min(top->lower_x, bottom->lower_x))) ||
        !(!(overlap(top->upper_x, top->lower_x,
        min(top->upper_x, bottom->upper_x),
        min(top->lower_x, bottom->lower_x))) ||
        return 0;

    }
    return 0;

    /* Test if some section fully inside the other section
    .....
    .....
    .....
    int full_in_section(first, second)
    {
        if ((!(first->left_x > second->right_x) || (first->right_x <= second->right_x) ||
            (first->upper_x > second->upper_x) || (first->lower_x <= second->lower_x)))
            return 1;
        else
            return 0;
    }

    /* Test if two section are neighbors (belong to same column)
    .....
    .....
    .....
    int same_section(section, first, sec, num)
    {
        struct section *sec;
        int first, sec, num;

        int x1, y1, x2, y2, i, j, counter;
        int in_section, in_column, in_section, in_column;
        struct section *sec;
        struct section *sec;
        int are_in_col, are_in_row, are_in_col, are_in_row;
        char p;

        printf("In same section index 1: %d index 2: %d\n", section->first->index,
            section->sec->index);

        if (section->first->left_x == section->sec->left_x)
        {
            top = section->first;
            bot = section->sec;
        }
        else
        {
            top = section->sec;
            bot = section->first;
        }

        if ((!(overlap(top->upper_x, top->lower_x,
            top->upper_x, top->lower_x) || 0)) return 0;

        if ((!(top->upper_x == top->lower_x) || (top->upper_x == 0)) return 0;
        if ((!(top->first->index == top->sec->index) ||
            if ((top->first->index == top->sec->index)
            {
                return 0;
            }
            else if ((top->upper_x == top->lower_x) ||
                return 0;
            }
    }

```

【☒ 1 1 1】

[illegible]

【図 1 1 2】

[illegible]

【例 1 1 4】

[illegible]

【☒ 1 1 5】

[illegible]

-126-

[illegible]

【图 1 1 9】

[illegible]

—129—

[illegible]

[illegible]

[illegible]

【図125】

```

/*****
***** File: block_util.c
***** Author: Shin-Yuan Wang
***** Date: 1-28-93
***** Copyright 1993 Canon Information Systems
*****/

/***** Test if covered range overlap *****/

int if_overlap(a1, a2, b1, b2)
{
    int a1, a2, b1, b2;
    if ((a2 < b1) || (b2 < a1)) return 0;
    return 1;
}

/***** minimum of two integers *****/

int min(a, b)
{
    int a, b;
    if (a <= b) return a;
    else return b;
}

/***** Ratio of overlapped *****/

float cover_range(a1, a2, b1, b2)
{
    int a1, a2, b1, b2;
    float range = 0.0;
    int x1, x2;
    if ((a2 < b1) || (b2 < a1)) return range;
    else {
        x1 = max(a1, b1);
        x2 = min(a2, b2);
    }
    range = ((float)(x2-x1-1)/(float)(a2-a1+1));
    return range;
}

```

【図131】

```

/*****
***** Filename: blockline.h
***** Author: Shin-Yuan Wang
***** Date: 1-15-93
***** Copyright 1993 Canon Information Systems
*****/

#define EMPTY 0
#define OVERLAP 1
#define NEXT 2
#define SURROUND 3
#define SURROUNDOUT 4

#define TH 1.7
#define WIDTHSTD 1.2
#define CAPLENGTH 0

#define DOT_BOTTOM 0.9
#define DOT_WIDTH 1
#define LDR/OVERLAP 0.1

#define LEFTPAD 4*10
#define RIGHTPAD 0*10
#define MINER_WD 3

typedef int kllcondition;

extern void ka_draw_box();
extern void convert_to_block();
extern void transfer_block();
extern void remove_cerblock();
extern int max();
extern int min();
extern float cover_range();
extern struct blockline **new_linepointer();
extern void sort_linepointer();

```

【図127】

```

/*****
*****      Filename: block_main.h
*****      Author:  Shin-Yuan Wang
*****      Date:    1-15-93
*****      Copyright 1992 Canon Information Systems
*****/

#define TESTTIME 20
#define TEXT_BLOCK 0x11;
#define LINEPICTURE_BLOCK 0x31
#define NPICTURE_BLOCK 0x31
#define TABLE_BLOCK 0x41
#define LINE_BLOCK 0x51
#define FRAME_BLOCK 0x61
#define UNKNOWN_BLOCK 0x71

#define TEXT_COLOR 0
#define TABLETEXT_COLOR 1
#define LINE_COLOR 2
#define PICTURE_COLOR 3
#define NPICTURE_COLOR 4
#define UNKNOWN_COLOR 5
#define TABLE_COLOR 6
#define FRAME_COLOR 7
#define LINECOLOR 8

enum {SP_READ = 0x01,
      SP_READUC = 0x02,
      SP_CURR = 0x04,
      SP_BLOCK = 0x08};

struct block_image
{
    int width;
    int height;
    unsigned char** pixel;
};

int koff(TESTTIME)[2] = {(0, 1), (1, 2), (0, 1), (1, 2), (0, 1),
                          (1, 2), (0, 0), (0, 0), (1, 1), (1, 1),
                          (2, 2), (2, 2), (1, 0), (2, 1), (1, 0),
                          (3, 1), (0, 1), (0, 1), (1, 2), (1, 2)};

int yoff(TESTTIME)[2] = {(0, 0), (0, 0), (1, 1), (1, 1), (2, 2),
                          (1, 1), (0, 1), (1, 1), (0, 1), (1, 1),
                          (0, 1), (0, 2), (0, 1), (0, 1), (1, 2),
                          (1, 1), (1, 2), (0, 1), (1, 2), (0, 1)};

#define TOTALCOLOR 7
char *color_label[]={
    "Text",
    "Text in the Table",
    "Picture",
    "Line Drawing Graphics",
    "Solid Line (Vertical or Horizontal)",
    "Table",
    "Frame"};

int color_code[]={TEXT_COLOR, TABLETEXT_COLOR, NPICTURE_COLOR, PICTURE_COLOR,
                  LINE_COLOR, TABLE_COLOR, FRAME_COLOR};

```

【図132】

```

/*****
*****      Filename: blocksection.h
*****      Author:  Shin-Yuan Wang
*****      Date:    1-15-93
*****      Copyright 1992 Canon Information Systems
*****/

#define GAP 16
#define OVERLAP 0.8
#define HOVERLAP 0.5
#define LEVELOVERLAP 0.8
#define LEFTSIDE 0x00
#define LEFTSIDE 0x00
#define LEFTTOPLONGER 0x00
#define LEFTBOTLONGER 0x00
#define RIGHTSIDE 0x01
#define RIGHTSIDE 0x01
#define RIGHTTOPLONGER 0x01
#define RIGHTBOTLONGER 0x01
#define TEXTPIXEL 0x00

#define NOISE_COLOR 8

extern int line_inside();
extern int max();
extern int min();
extern int is_overlap();
extern void set_attr_box();
extern float cover_range();
void exit();

struct segment_gap
{
    int column_no;
    int *begin_column;
    int *end_column;
};

```


—137—

[illegible]

【図135】

```

/*****
***** Filename: Table.h *****
***** Author: Shin-Yuan Wang *****
***** Date: 1-15-92 *****
***** Copyright 1992 Canon Information Systems *****
*****/

#include "blockgeneral.h"

#define NAME 0
#define OFFSET 1
#define ROW 1
#define OLD 1
#define ROWSTOP 0
#define STOP 1
#define TABLE_LENGTH 1

struct white_for_table
{
    int status;
    struct whiteblock *white;
    struct white_for_table *previous, *next;
};

```

【図136】

```

/*****
***** Filename: Images.h *****
***** Author: Shin-Yuan Wang *****
***** Date: 1-15-92 *****
***** Copyright 1992 Canon Information Systems *****
*****/

#include <stdio.h>
#include <stdlib.h>

struct images
{
    int size;
    int size2;
    int dot;
    unsigned char *pixel;
};

```

【図142】

```

/*
* Filename: autoinput.h
* Header file for autoinput.c
*/

/* Function Definitions for mymalloc.c */
#ifdef CPP
/* Definition for C++ files */
extern "C" {
    void auto_initialize();
    int auto_input_int();
    void auto_terminate();
}
#else
/* Definition for C files */
void auto_initialize();
int auto_input_int();
void auto_terminate();
#endif

```

【図145】

```

/*****
void herovpp_terminate(mnlist)
    struct MoniterovppList *mnlist; /* The list */
{
    free_vector(mnlist->size); /* Deallocate memory */
    mnlist->size = 0; /* Set so won't be accidentally used */
}

/*****
void herovpp_add(mnlist, a, b)
    struct MoniterovppList *mnlist; /* The list */
    int a, b; /* Start and stop of the herovpp */
{
    if (mnlist->size == mnlist->allocated) {
        printf("MoniterovppList FULL!!\n");
        exit(1);
    }
    mnlist->list[mnlist->size++] = a;
    mnlist->list[mnlist->size++] = b;
}

```

【図244】

```

colors[x].flag = Dored; Dorgreen; Doblue;

colors[x].red = the_color;
colors[x].green = the_color;
colors[x].blue = the_color;
the_color += 50;

}
XQueryColors(d, del_colormap, colors, ncolors);

my_cmap = XCreateColormap(d, DefaultRootWindow(d), DefaultVisual(d, scr),
                          AllocAll);
XStoreColors(d, my_cmap, colors, ncolors);
}

```


—140—

【图 140】

-141-

【図141】

```

/* filename: autoinput.h
 * This file contains the code to input from a filename
 * .....
 */
#include <stdio.h>
#include <stdlib.h>
/* Global variables */
int auto_eof_detected,
FILE *auto_fp;
/* The file pointer */
int save_count;
/* Making the input to a file */
/* Making the input before newline */
char filename[100];
/* The filename */
.....
void auto_initialize()
{
    int not_eof = 1;
    auto_eof_detected = 0;
    while(not_eof)
    {
        printf("Input filename for auto-input: ");
        scanf("%s", filename);
        auto_fp = fopen(filename, "r");
        if (auto_fp == NULL) /* Read open failed */
        {
            auto_fp = fopen(filename, "w");
            if (auto_fp == NULL) /* Write open failed */
            {
                printf("Can't open file (%s)", filename);
                printf("Enter 1 at the next prompt.\n");
                save_count = 1;
                auto_eof_detected = 1;
                not_eof = 0;
            }
            else {
                printf("Opening existing file (%s)", filename);
                not_eof = 0;
            }
        }
        printf("Do you wish to save the auto-input into this file (1 = yes)? ");
        save_count = auto_input_int();
        printf("Enter 1 at the next prompt.\n");
        save_count = 0;
    }
}

int auto_input_int()
{
    int i;
    if (auto_eof_detected)
        printf("\n(Warning) Input (-) = newline, -9999 = exit) : ");
}

```

```

scanf("%d", &i);
if (i == -9999)
{
    fclose(auto_fp);
    exit(0);
}
if (save_count == 1)
{
    /* Print new lines */
    {
        printf(auto_fp, "\n"); /* Print the new lines */
        printf(auto_fp, "\n");
        fflush(auto_fp);
        save_count = 0;
    }
    return auto_input_int(); /* And input again */
}
printf(auto_fp, "%d ", i);
fflush(auto_fp);
if (save_count > 20)
{
    save_count = 0;
    printf(auto_fp, "\n");
    fflush(auto_fp);
    return 1;
}
if (feof(auto_fp, "r", &i) == EOF)
{
    printf("\n..... END OF FILE DETECTED .....");
    if (save_count == 1)
    {
        fclose(auto_fp);
        printf("Can't open file (%s)", filename);
        if (auto_fp == NULL)
        {
            printf("..... WARNING: CANNOT APPEND TO FILE. Is %s", filename);
            save_count = 0;
        }
        printf(auto_fp, "\n\n");
        auto_eof_detected = 1;
        return auto_input_int();
    }
    printf("\n\n");
    return 1;
}
void auto_eof_detected()
{
    printf(".....");
}

```


【図146】

```

/* filename: charlist.h
 * Header file for charlist.c
 */
.....

/* ENUM: CharType: Different types of characters */
typedef enum {
    CHARACTER_NULL,
    CHARACTER_UPPERCASE,
    CHARACTER_LOWERCASE,
    CHARACTER_PUNCT,
    CHARACTER_NON_PRINTED,
    CHARACTER_OTHER
} CharType;

typedef struct {
    int left;
    int right;
    int curved;
    int curved;
    int 'curved';
    int 'curved';
    struct Character *next;
    struct Character *prev;
} CharList;

/* Contains information about an undefined character in a page
 * Recognition hasn't been tried on it */
.....
/* The type of character */
/* The leftmost pixel in this character */
/* The rightmost pixel in this character */
/* The left side of the character (inclusive)
 * (for good luck) */
/* The right side of the character (exclusive)
 * (for good luck) */
/* The left side of the character (exclusive)
 * (for good luck) */
/* The right side of the character (inclusive)
 * (for good luck) */
/* The left side of the character (exclusive)
 * (for good luck) */
/* The right side of the character (inclusive)
 * (for good luck) */
/* The left side of the character (exclusive)
 * (for good luck) */
/* The right side of the character (inclusive)
 * (for good luck) */
.....
/* Struct Character
 * Contains information about the list of characters */
struct Character {
    CharType type;
    CharList *list;
    int left;
    int right;
    int curved;
    int 'curved';
    struct Character *next;
    struct Character *prev;
};

/* Contains information about the list of characters */
struct CharList {
    struct Character *chr;
    int size;
    int allocated;
    int vector_size;
    struct Character *first;
};

/* Contains information about a list for non-zero upper */
.....
/* Zero 'pp' lists stuff
 * Contains information about a list for non-zero upper */
.....

```

```

.....
/* Array: In a state, 2nd - stop */
/* Memory allocated (2, 4, 8, ...) */
/* How many used (2, 4, 8, ...) */
.....
/* Define RotLines
 * .....
```


【図148】

```

// Copy clipped image
for (j=0; j<2; j++)
    temp_image[(1-11)/2-C1] = pixel[(11)/2];
allocate(temp_image, temp_imageC1); // Copy temp image to current image
for (j=0; j<2; j++)
    pixel[(11)/2] = temp_image[(11)/2];
}

// Read in part of a Tiff file
void chage::read_tiff(char *filename)
{
    FILE *tagfile; // Input file
    long int tagfile_position; // The our file position (in bytes)
    long int width; // Width of picture
    long int length; // Length of picture
    long int black; // Value of black pixel
    long int hps; // Bytes per strip
    long int start_offset; // Start of image data
    long int nstrips; // Number of strips
    long int type; // Tag type
    long int value; // Tag value

    if ((tagfile = fopen(filename, "rb")) == NULL)
        cout << "Unable to open file " << filename << "\n";
    exit(1);

    tagfile_position = 0;

    // Read in header and make sure it's local
    if (!getbytes(tagfile, tagfile_position,
        if (1 != 0x002a))
        cout << "Bad magic number for TIFF file\n";
        exit(1);

    // Find first file directory
    tagfile_position = 0;
    if (!getbytes(tagfile, tagfile_position,
        if (1 != 0x002a))
        cout << "Bad magic number for TIFF file\n";
        exit(1);

    // Tags in directory
    nstrips = getbytes(tagfile, tagfile_position);
    while (1 - nstrips > 0) {
        tag = getbytes(tagfile, tagfile_position);
        type = getbytes(tagfile, tagfile_position);
        value = getbytes(tagfile, tagfile_position);
        value = getbytes(tagfile, tagfile_position);
        if (1 != 11 44 (tag & 32768) == 0)
            cout << "Error: Tag length is 1\n";
    }
}

```

```

// Make the new image into size adjuster:
for (int i=0; i<2; i++)
    pixel[(11)/2] = value;
}

// Copy a section of a image into another. Image = reference image
// Range is [R1..R2], [C1..C2]
void chage::clip_image(image, int R1, int R2, int C1, int C2)
{
    int r, c;

    if (R1 < R2)
        if (C1 < C2)
            // Swap R1 & R2
            // Swap C1 & C2
            // Check limits
            if (R1 < 0) R1 = 0;
            if (R2 > 1) R2 = 1;
            if (C1 < 0) C1 = 0;
            if (C2 > 1) C2 = 1;
            if (R1 > R2) R1 = R2;
            if (R2 > 1) R2 = 1;
            if (C1 > C2) C1 = C2;
            if (C2 > 1) C2 = 1;

    allocate(R2-R1, C2-C1);

    for (r=R1; r<R2; r++)
        for (c=C1; c<C2; c++)
            pixel[(r-C1)] = image_pixel[(r-11)][(c-C1)];
}

// Clipper: reduce the size of an image by removing the blank
// edges. It will adjust the image file.
void chage::clip_image()
{
    int R1, R2, C1, C2; // The starting and stopping rows of new image
    int i, j; // Temporary image
    R1 = 0; // Initialise to extremes
    R2 = 0;
    C1 = 0;
    C2 = 0;

    for (i=0; i<imageC1; i++)
        for (j=0; j<imageC2; j++)
            if (pixel[(11)/2] != 0)
                if (i < R1) R1 = i;
                if (i > R2) R2 = i;
                if (j < C1) C1 = j;
                if (j > C2) C2 = j;
    }

    if (R1 > R2) R1 = R2 + 0;
    if (C1 > C2) C1 = C2 + 0;

    temp_allocate(R2-R1+1, C2-C1+1); // New image is this size
}

```

【図149】

```

    } // while
    while(tagfile_position < source_offset)
    {
        l = getbyte(tagfile, tagfile_position);
    }
    //----- READ IN FILE -----
    allocate(int) length, (int) width;
    int vi;
    for (i = 0; i < length; i++)
    {
        for (j = 0; j < width; j++)
        {
            if ((j % 8) == 0)
            {
                v = getbyte(tagfile);
                tagfile_position++;
                if (v == EOF)
                {
                    cout << "Error: Premature End of File found in file filein";
                    exit(1);
                }
            }
            if ((i % 128) == black-128)
            {
                pixel[i][j] = 0;
                pixel[i][j+1] = 1;
                v = v << 1;
            }
        }
        fclose(tagfile);
    }
    //----- READ IN FILE -----
    // Read in a byte. File position is updated
    // and the value is stored in the variable
    int change = getbyte(tagfile);
    long int file_position, int eoferror;
    int ci;
    ci = getc(f);
    if ((ci == EOF) || (eoferror == 1))
    {
        cout << "Error: Premature End of File (in getbyte())";
        exit(1);
    }
    // File position++
    return ci;
}
// Read in an integer from a file. (LWS, LWS)
long int change = getbyte(FILE *f, long int ci)
{
    int ci, ci2;
    ci = getbyte(f, f);
    ci2 = getbyte(f, f);
    return ci * 256 + ci2;
}
// Read in a long integer from a file. (LWS, LWS)
long int change = getbyte(FILE *f, long int ci)
{
    int ci, ci2, ci3, ci4;
    ci = getbyte(f, f);
    ci2 = getbyte(f, f);
    ci3 = getbyte(f, f);
    ci4 = getbyte(f, f);
    return ci * 16777216 + ci2 * 65536 + ci3 * 256 + ci4;
}
}
exit(1);
}
switch (tag)
{
    case 0x00FF:
        if (value != 0)
        {
            cout << "Can't deal with non-standard NewRufFileType";
            exit(1);
        }
        break;
    case 0x007F:
        if (value != 1)
        {
            cout << "RufFile type not full resolution";
            exit(1);
        }
        break;
    case 0x0100:
        width = value;
        break;
    case 0x0101:
        length = value;
        break;
    case 0x0102:
        if (value != 1)
        {
            cout << "Error: Bits/sample must be 1";
            exit(1);
        }
        break;
    case 0x0103:
        if (value != 1)
        {
            cout << "Compression Used Error";
            exit(1);
        }
        break;
    case 0x0104:
        black = value;
        break;
    case 0x0105:
        if (value < 1) // Threshold tag
        {
            cout << "Threshold tag < literal values";
            exit(1);
        }
        break;
    case 0x0111:
        offset = value; // Offset to Marker Data
        break;
    case 0x0117:
        type = value;
        break;
    case 0x0118:
        break;
    case 0x0119:
        break;
    case 0x0121:
        break;
    default:
        if (value > 255) // If (value > 255) == 0)
        {
            cout << "Error: Number is unsupported tag appears in the TIF file";
            break;
        }
        // switch
}

```

【图 167】

```

return void cut_dash(cdashget_line, struct Charlist* clist);
cut_dash((*line_to_process), clist);
#endif

if(defs COMBINE)
    charlist_combine((*line_to_process), clist);
endif

if(defs PRINT_FINAL_CHARLIST)
    ////////// PRINT the charaterlist
    ma_clr_win();
    print_charlist((*line_to_process), clist, 10, 10);
    ma_flush();
endif

if(defs DISPLAY)
    ma_flush();
endif
printf("Displaying Final Segmentation. Enter 0 to continue ");
int chany; int chany;
chany = auto_input_int();

return terminate_to_nextterminaline; // == DELETES ==

```

【図 220】

```

/*
 * Filename: autoinput.h
 * Header file for autoinput.c
 *
 * .....
 */

/* Function Definitions for mymalloc.c */
#ifdef CPP
/* Definition for C++ files */
extern "C" {
    void auto_initialize();
    int auto_input_line();
    void auto_terminate();
}
#else
/* Definition for C files */
void auto_initialize();
int auto_input_line();
void auto_terminate();
#endif

```

【图 151】

```

/*
 * Filename: mymalloc.h
 * Header file for mymalloc.c
 *.....
 */

/* Function Definitions for mymalloc.c */
#ifdef CPP
/* Definition for C++ files */
extern "C" {
    int *ivector(int n);
    void free_ivector(int* v);
    float *fvector(int n);
    void free_fvector(float* v);
    int *ivector_ranged(int n, int m);
    void free_ivector_ranged(int* v, int n);
    int **imatrix(int r, int c);
    void free_imatrix(int* m, int r);
    char **cmatrix(int r, int c);
    void free_cmatrix(char** m, int r);
}
#else
/* Definition for C files */
/* The header file for malloc - */
int *ivector();
void free_ivector();
int *fvector();
void free_fvector();
int *ivector_ranged();
void free_ivector_ranged();
int **imatrix();
void free_imatrix();
char **cmatrix();
void free_cmatrix();
#endif

```

【图 2 3 5】

[1 5 3]

```

for (i=c1->bot - 4; i < c2->bot; i++)
    botpos += c2->curves[i] - c1->curves[i];
}
// Compute position and exit
*d3 = botpos - toppos;
if (i==(float)*d3) / ((float) width_small);
if (i < PASC_VALID_DIST)
    return 1;
else
    return 0;
}

// This routine is called from comb_dier() with a pair of characters
// that probably isn't an i or j.
// Character. This determines if close enough to be combined.
int comb_dier_jell()
{
    struct Character *c1; // The first character
    struct Character *c2; // The second character
    int i; // Which character is higher
    int *d3; // The last value
    int width_small; // The width of the smaller character

    int toppos, botpos; // Positions of the top and bottom
    float f;

    if (i33 == 1)
        // Case when c1 is higher than c2
        // Find bottom pos of c1
        if ((c1->bot - c1->top) < 4)
            botpos = c1->bot;
        else
            botpos = 0;
        for (i=c1->bot - 4; i < c1->bot; i++)
            botpos += c1->curves[i] - c1->curves[i];
    }
    // Find top pos of c2
    if ((c2->bot - c2->top) < 4)
        // If character very small, use .loc
        else
            // If character very small, use .loc
            // Compute position and exit
            *d3 = toppos - botpos;
            f = ((float) *d3) / ((float) width_small);
            if (i < PASC_VALID_DIST)
                return 1;
            else
                return 0;
    }
    // c2 is higher than c1
    // Find top pos of c1
    if ((c1->bot - c1->top) < 4)
        // If character very small, use .loc
        else
            // If character very small, use .loc
            // Compute position and exit
            *d3 = toppos - botpos;
            f = ((float) *d3) / ((float) width_small);
            if (i < PASC_VALID_DIST)
                return 1;
            else
                return 0;
    }
    // c2 is higher than c1
    // Find bottom pos of c2
    if ((c2->bot - c2->top) < 4)
        // If character very small, use .loc
        else
            // If character very small, use .loc
            // Compute position and exit
            *d3 = toppos - botpos;
            f = ((float) *d3) / ((float) width_small);
            if (i < PASC_VALID_DIST)
                return 1;
            else
                return 0;
    }
}

```


【図154】

```

1  if ((height_c1 < FONT_HEIGHT) && height_c2)
2  return comb_dsize(c1, c2, h1, d1, width_small);
3  return comb_dsize(c1, c2, h1, d1, width_small);
4  }
5  // Else c1 is higher than c2
6  if ((height_c1 < FONT_HEIGHT) && height_c2)
7  return comb_dsize(c1, c2, h1, d1, width_small);
8  else
9  return comb_dsize(c1, c2, h1, d1, width_small);
10 }
11
12 // This routine uses if larger characters are overlapping enough (1st: end
13 // right side.) used for identifying type 1 percent alone.
14 // Returns 0, if overlapping less than 1 percent, 0 otherwise
15 int comb_dsize_overlap(
16     struct character *c1, // The first character
17     struct character *c2) // The second character
18 {
19     int i1, i2; // The length of the characters
20     int j1, j2; // The length of the characters
21     struct character *swap; // Swap the characters
22     int swap1; // Swap the characters
23     int swap2; // Swap the characters
24     i1 = c1->right - c1->left; // Compute width
25     i2 = c2->right - c2->left; // Compute width
26     if (i1 > i2) // c1 is longer
27     {
28         swap = c1;
29         c1 = c2;
30         c2 = swap;
31         swap1 = i1;
32         i1 = i2;
33         i2 = swap1;
34     }
35     if ((float) i1 / (float) i2 < FRCDEF_TYH1_MIN)
36     return 0;
37     if ((float) i1 / (float) i2 > FRCDEF_TYH1_MAX)
38     return 0;
39     if (c2->left < c1->left)
40     {
41         swap = c1;
42         c1 = c2;
43         c2 = swap;
44     }
45     if (c2->right < c1->right)
46     {
47         swap = c1;
48         c1 = c2;
49         c2 = swap;
50     }
51     if ((float) (c2 - c1) / (float) i1 > FRCDEF_TYH1_MAX)
52     return 1;
53     else
54     return 0;
55 }
56
57 // Find bottom pos of c2
58 if (c1->bot - c2->bot < 4)
59 botpos = c2->bot; // If character very small, use .loc
60 else
61 botpos = 0;
62
63 // Compute position and exit
64 for (i = 0; i < c1->bot; i++)
65 botpos = c2->bot;
66
67 // Compute position and exit
68 if (c1->bot < c2->bot)
69 return 1;
70 if (i < FRCDEF_TYH1_MIN)
71 return 1;
72 return 0;
73 }
74
75 // This routine returns 1 if the two characters are overlapping
76 // to bottom of higher one. Returns 0 - bad, 1 - good. Also updates d1
77 // Position is generated to be relative to dot height
78 int comb_dsize(
79     struct character *c1, // The first character
80     struct character *c2) // The second character
81 {
82     int h1, h2; // Which character is higher
83     int d1, d2; // The dot value
84     int width_c1; // The size of c1 (width)
85     int width_c2; // The size of c2
86     int width_small; // The size of the smaller character
87     int height_c1; // The height of c1
88     int height_c2; // The height of c2
89     float f;
90     width_c1 = c1->right - c1->left;
91     width_c2 = c2->right - c2->left;
92     if (width_c1 < width_c2)
93     width_small = width_c1;
94     else
95     width_small = width_c2;
96     if (c1->bot < c2->bot) // If it's been computed, don't recompute it
97     {
98         f = ((float) c1->bot) / (float) width_small;
99         if (f < FRCDEF_TYH1_MIN)
100         return 1;
101         else
102         return 0;
103     }
104     height_c1 = c1->bot - c1->top;
105     height_c2 = c2->bot - c2->top;
106     if (h1 < h2) // c1 is higher than c2

```


[illegible]

[illegible]

【図160】

```

// Check for the dash between a character, and break it if needed
void cd_char(
    char** line, // The line
    struct Charlist *clist, // Pointer to charlist
    struct Character *ch, // Pointer to the character we're on (modified
                          // to skip past newly inserted divisions)
    int *hist) // The histogram (previously allocated)
{
    int p;

    // Reject if pixel width is too small.
    if ((ch->right - (ch->left) < MIN_ABS_WIDTH_TO_CONSIDER)
        return;

#ifdef PRINT_INFO
    xs_out_idb();
    print_character(line, "ch", 10, 10);
    xs_flush();

    print(("Character info:\n");
    print((" top/bot = (%d, %d)\n", (ch->top, (ch->bot);
    print((" left/right = (%d, %d)\n", (ch->left, (ch->right);
#endif

    // Compute the histogram
    hist_histogram(line, ch, hist);

    // Break left side?
    p = cd_left_dash_cut(line, ch, hist);
    if (p != 0)
    {
        cd_cut_character(line, clist, ch, p); // Cut the character
#ifdef PRINT_INFO
        print(("Left dash cut\n");
        print(("Character info:\n");
        print((" top/bot = (%d, %d)\n", (ch->top, (ch->bot);
        print((" left/right = (%d, %d)\n", (ch->left, (ch->right);
#endif
    }

    // Break right side?
    p = cd_right_dash_cut(line, ch, hist);
    if (p != 0)
    {
        cd_cut_character(line, clist, ch, p); // Cut the character
#ifdef PRINT_INFO
        print(("Right dash cut\n");
#endif
    }
}

//
//.....
// MAIN ROUTINE
//.....
//
// This will cut the dashes of of characters
void cut_dash_ranges(
    char** line, // The image
    struct Charlist *clist, // The current charlist
    int *histogram, // The area of the histogram
    struct Character *ch, // The character we're processing

```

【図195】

```

/*
 * Filename: myalloc.h
 * Header file for myalloc.c
 */

/*
 * Function definitions for myalloc.c */
#ifdef C
/* Definition for C++ files */
extern "C" {
    int *vector(int n);
    void free_vector(int* v);
    float *vector(float n);
    void free_vector(float* v);
    int *vector_ranged(int n, int a);
    void free_vector_ranged(int* v, int a);
    int **matrix(int x, int y);
    void free_matrix(int** m, int x);
    char **matrix(int x, int y);
    void free_matrix(char** m, int x);
}
#else
/* Definition for C files */
/* The header file for malloc */
int *vector();
void free_vector();
float *vector();
void free_vector();
int **matrix();
void free_matrix();
char **matrix();
void free_matrix();
#endif

```

【図223】

```

/*.....*/
void seroyp_add(intlist, a, b)
    struct seroyp_list *nlist; /* The list */
{
    free_vector(nlist->list); /* Deallocate memory */
    nlist->allocated = 0; /* Set so won't be accidentally used */
    nlist->size = 0;
}

/*.....*/
void seroyp_add(intlist, a, b)
    struct seroyp_list *nlist; /* The list */
    int a, b; /* Start and stop of the seroyp */
{
    if (nlist->size+1 > nlist->allocated) {
        print(("Warning: list FULL!\n");
        exit(1);
    }
    nlist->list[nlist->size++] = a;
    nlist->list[nlist->size++] = b;
}

```

[図161]

```

loc = loc_count + 0; // for position of the char (on the fly)
for (z=0; z<w; z++)
{
    while (curvel[z] < curvel[z]) // loop until they're equal
    { (libe.pixel[z] - curvel[z]) < 1; } // if there is a non empty space
    // then exit (left border not needed)
    break;
    // There's an empty space to the right
    curvel[z] += 1;
    while (curvel[z] < curvel[z]) // slide the border over to meet it
    { (libe.pixel[z] - curvel[z]) < 1; } // Do the same with the right border
    else
    {
        curvel[z] += 1;
    }
    if (curvel[z] != curvel[z]) // there is a pixel on this line...
    {
        loc = curvel[z] + curvel[z];
        loc_count++;
    }
    if (loc_count == 0)
    {
        loc = loc + 4 / loc_count; // computes the position (used for .loc;
        pos = loc/4; // compute the real position
    }
    // Find left and right
    left = curvel[top];
    right = curvel[top];
    for (z=top; z<w; z++)
    {
        if (curvel[z] < left)
        {
            left = curvel[z];
        }
        if (curvel[z] > right)
        {
            right = curvel[z];
        }
    }
    if (libe == 0)
    {
        loc = 4 * (left+right)/2;
        pos = (left+right)/2;
    }
    for (z=top; z<w; z++) // fill in the blank lines with pos
    {
        curvel[z] = curvel[z] + pos;
    }
    // Copy all the new attributes in
    (*ch)-libe = loc;
    (*ch)-top = top;
    (*ch)-bot = bot;
    (*ch)-left = left;
    (*ch)-right = right;
}

// This routine will cut with libe the
// and *ch returns as the 'ch' pointer to the right half of the character
// changed line
struct charlist *cllc;
struct Character *ch;
// The character list
// The character to be cut
// Have to cut the character
int y;

```

```

// filename: level1.c
// This file contains the routines to do level 1 segmentation (slightly
// modified) using method 1 (libe's approach)
//
//
//
// The libe
//
// include charlist.h
// include climp.h
// include normalloc.h
//
//
// The constants
#define MAX_CUT_THRESHOLD 2
#define MAX_PIXEL_THRESHOLD 16
//
// This routine will adjust left, right, curvel, curvel, top, bottom
// and pos to represent a new character in the character list, then it
// calls charlist_replace to replace the new character
// Possible optimization: usually, this is called when a character is
// void libe_replace(charlist *cllc, int x, int y, int w, int h)
//
// changed line, // pointer to the character we're on
// struct character *ch //
// int top, bot, left, right; // The new left and right stuff.
// int pos; // pos = position, loc = pos + 4
// int loc, loc_count, pos; // Northward, set to (*ch)-curvel and 2
// int flag;
// curvel = (*ch)-curvel;
// curvel = (*ch)-curvel;
//
// Adjust the top of the character, if it needs to move down
// flag = 0; // the found when flag == 1,
// for (top = (*ch)-top; top < (*ch)-bot; top++)
// {
//     if (libe.pixel[top] < 1) // stop at the first on pixel
//     {
//         flag = 1;
//         break;
//     }
// }
//
// Adjust the bottom of the character
// flag = 0; // not found when flag == 1,
// for (bot = (*ch)-bot-1; bot > (*ch)-top; bot--)
// {
//     if (libe.pixel[bot] < 1) // stop at the first on pixel
//     {
//         flag = 1;
//         break;
//     }
// }
//
// bot--; // (like bot exclusive)
//
// Now adjust curvel and curvel, update left and right to 'shrink'
// slightly around the border of the character (move the borders up to the
// // edge of the characters.

```


—160—

[illegible]

—164—

[illegible]

【図170】

```

int i;
int dummy;

// The cursor in the projection
for (i=1; i<proj_size; i++) // For each border between points i-1 & i
{
    if (i[proj[i]-1] == 0) // If right border
        dummy = breakline_and_break(blister, (i-1)*proj_size, blister->vector_height, blister->size-1);
    else if (i[proj[i]-1] == 1) // If left border
        dummy = breakline_and_break(blister, i*proj_size, blister->vector_height, blister->size, 0);
    else if (i[proj[i]-1] == 2) // If left edge
        dummy = breakline_and_break(blister, i*proj_size, blister->vector_height, blister->size, 0);
}

// Find minimum value in a vector over a specified range
void find_minimum_in_vector(
    int a[],
    int n,
    int min,
    int max)
{
    // The input vector
    // The range to search over (inclusive)
    // Output: minimum value
    // Output: it's index
    // Set first value as default
    min = a[1];
    int i;
    for (i=2; i<=n; i++)
        if (a[i] < min)
            min = a[i];
}

// Find minimum in a valley for a rotated profile
void find_minimum_in_valley(
    int a[],
    int n,
    int min,
    int max)
{
    // The image
    // The breakline for the image
    // The range of breaks to search through
    // Where to break on the right
    // The angle of rotation to compute
    // The minimum value -1 if no good cross
    // Where the minimum is located at (-1 = bad)
    // Range to look for
    // The rotated projection
    // Info on the projection
    // The minimum value of the projection
    int i, j;
    int a, allocated_a;
    int max_a, index_max;
    int min_a;

    // Allocate memory for a[]
    allocated_a = image_size * sizeof(int); // This should be enough
    a = (int*) malloc(allocated_a);

    // Find the rotated projection
    find_projection(blister, a, blister->size, blister->size, 0);

    // Find maximum projection
    max_a = a[1];
    for (i=2; i<=n; i++)
        if (a[i] > max_a)
            max_a = a[i];
}

```

```

// Find maximum projection
max_value = 0;
for (i=0; i<max_x; i++)
    if (a[i] > max_value)
        max_value = a[i];

// Choose range to ignore points before first peaks on left and right
i1 = 0;
i2 = max_x-1;

if (i1 == 0)
    // Find left peak while slope is still rising, and above its contour height
    while ((a[i1+1] < a[i1]) || (a[i1] < a[i1+2])) { i1++; }
    a[i1] = MAX_CONTOUR_HEIGHT + max_value;
    if (i1 == 0) // If got to left edge
        min = min_at = -1;
    else // Return no minimum found
        return;
}

// Find right peak while slope is falling (from right to left)
while ((a[i2-1] < a[i2]) || (a[i2] < a[i2-2])) { i2--; }
a[i2] = MAX_CONTOUR_HEIGHT + max_value;

if (i1 == i2) // If reached the left point
    min = min_at = -1;
else // Return no minimum found
    return;
}

// Find minimum in vector(a, i1, i2, min, min_at) // Find the minimum
return min;
}

// The above is mainly copied from void find_min_valley_at_index ...
// The following section is VERY VERY poorly written!!!!!!
for (i=cursor_col_left; i<cursor_col_right-1; i++) // Draw line above
    cursor_x = i;

// Plot the image
printf("image with breaks (image, list, image), image, cursor_x, cursor_y);
return break; // Find position of the c line and plot it
// Find first break
int c = cursor_x - MAX_IMAGE_ALIGN + c + 1;
for (i=cursor_x - image_size + 1; i<cursor_x + image_size + 1; i++)
    plot(i, a[cursor_y], c, then, cursor_x - 4, cy + 5); // Right line
cursor_x = image_size + 5;

// Plot the histogram
int i1 = 0;
int i2 = 0;
int i3 = 0;
int i4 = 0;
int i5 = 0;
int i6 = 0;
int i7 = 0;
int i8 = 0;
int i9 = 0;
int i10 = 0;
int i11 = 0;
int i12 = 0;
int i13 = 0;
int i14 = 0;
int i15 = 0;
int i16 = 0;
int i17 = 0;
int i18 = 0;
int i19 = 0;
int i20 = 0;
int i21 = 0;
int i22 = 0;
int i23 = 0;
int i24 = 0;
int i25 = 0;
int i26 = 0;
int i27 = 0;
int i28 = 0;
int i29 = 0;
int i30 = 0;
int i31 = 0;
int i32 = 0;
int i33 = 0;
int i34 = 0;
int i35 = 0;
int i36 = 0;
int i37 = 0;
int i38 = 0;
int i39 = 0;
int i40 = 0;
int i41 = 0;
int i42 = 0;
int i43 = 0;
int i44 = 0;
int i45 = 0;
int i46 = 0;
int i47 = 0;
int i48 = 0;
int i49 = 0;
int i50 = 0;
int i51 = 0;
int i52 = 0;
int i53 = 0;
int i54 = 0;
int i55 = 0;
int i56 = 0;
int i57 = 0;
int i58 = 0;
int i59 = 0;
int i60 = 0;
int i61 = 0;
int i62 = 0;
int i63 = 0;
int i64 = 0;
int i65 = 0;
int i66 = 0;
int i67 = 0;
int i68 = 0;
int i69 = 0;
int i70 = 0;
int i71 = 0;
int i72 = 0;
int i73 = 0;
int i74 = 0;
int i75 = 0;
int i76 = 0;
int i77 = 0;
int i78 = 0;
int i79 = 0;
int i80 = 0;
int i81 = 0;
int i82 = 0;
int i83 = 0;
int i84 = 0;
int i85 = 0;
int i86 = 0;
int i87 = 0;
int i88 = 0;
int i89 = 0;
int i90 = 0;
int i91 = 0;
int i92 = 0;
int i93 = 0;
int i94 = 0;
int i95 = 0;
int i96 = 0;
int i97 = 0;
int i98 = 0;
int i99 = 0;
int i100 = 0;
int i101 = 0;
int i102 = 0;
int i103 = 0;
int i104 = 0;
int i105 = 0;
int i106 = 0;
int i107 = 0;
int i108 = 0;
int i109 = 0;
int i110 = 0;
int i111 = 0;
int i112 = 0;
int i113 = 0;
int i114 = 0;
int i115 = 0;
int i116 = 0;
int i117 = 0;
int i118 = 0;
int i119 = 0;
int i120 = 0;
int i121 = 0;
int i122 = 0;
int i123 = 0;
int i124 = 0;
int i125 = 0;
int i126 = 0;
int i127 = 0;
int i128 = 0;
int i129 = 0;
int i130 = 0;
int i131 = 0;
int i132 = 0;
int i133 = 0;
int i134 = 0;
int i135 = 0;
int i136 = 0;
int i137 = 0;
int i138 = 0;
int i139 = 0;
int i140 = 0;
int i141 = 0;
int i142 = 0;
int i143 = 0;
int i144 = 0;
int i145 = 0;
int i146 = 0;
int i147 = 0;
int i148 = 0;
int i149 = 0;
int i150 = 0;
int i151 = 0;
int i152 = 0;
int i153 = 0;
int i154 = 0;
int i155 = 0;
int i156 = 0;
int i157 = 0;
int i158 = 0;
int i159 = 0;
int i160 = 0;
int i161 = 0;
int i162 = 0;
int i163 = 0;
int i164 = 0;
int i165 = 0;
int i166 = 0;
int i167 = 0;
int i168 = 0;
int i169 = 0;
int i170 = 0;
int i171 = 0;
int i172 = 0;
int i173 = 0;
int i174 = 0;
int i175 = 0;
int i176 = 0;
int i177 = 0;
int i178 = 0;
int i179 = 0;
int i180 = 0;
int i181 = 0;
int i182 = 0;
int i183 = 0;
int i184 = 0;
int i185 = 0;
int i186 = 0;
int i187 = 0;
int i188 = 0;
int i189 = 0;
int i190 = 0;
int i191 = 0;
int i192 = 0;
int i193 = 0;
int i194 = 0;
int i195 = 0;
int i196 = 0;
int i197 = 0;
int i198 = 0;
int i199 = 0;
int i200 = 0;
int i201 = 0;
int i202 = 0;
int i203 = 0;
int i204 = 0;
int i205 = 0;
int i206 = 0;
int i207 = 0;
int i208 = 0;
int i209 = 0;
int i210 = 0;
int i211 = 0;
int i212 = 0;
int i213 = 0;
int i214 = 0;
int i215 = 0;
int i216 = 0;
int i217 = 0;
int i218 = 0;
int i219 = 0;
int i220 = 0;
int i221 = 0;
int i222 = 0;
int i223 = 0;
int i224 = 0;
int i225 = 0;
int i226 = 0;
int i227 = 0;
int i228 = 0;
int i229 = 0;
int i230 = 0;
int i231 = 0;
int i232 = 0;
int i233 = 0;
int i234 = 0;
int i235 = 0;
int i236 = 0;
int i237 = 0;
int i238 = 0;
int i239 = 0;
int i240 = 0;
int i241 = 0;
int i242 = 0;
int i243 = 0;
int i244 = 0;
int i245 = 0;
int i246 = 0;
int i247 = 0;
int i248 = 0;
int i249 = 0;
int i250 = 0;
int i251 = 0;
int i252 = 0;
int i253 = 0;
int i254 = 0;
int i255 = 0;
int i256 = 0;
int i257 = 0;
int i258 = 0;
int i259 = 0;
int i260 = 0;
int i261 = 0;
int i262 = 0;
int i263 = 0;
int i264 = 0;
int i265 = 0;
int i266 = 0;
int i267 = 0;
int i268 = 0;
int i269 = 0;
int i270 = 0;
int i271 = 0;
int i272 = 0;
int i273 = 0;
int i274 = 0;
int i275 = 0;
int i276 = 0;
int i277 = 0;
int i278 = 0;
int i279 = 0;
int i280 = 0;
int i281 = 0;
int i282 = 0;
int i283 = 0;
int i284 = 0;
int i285 = 0;
int i286 = 0;
int i287 = 0;
int i288 = 0;
int i289 = 0;
int i290 = 0;
int i291 = 0;
int i292 = 0;
int i293 = 0;
int i294 = 0;
int i295 = 0;
int i296 = 0;
int i297 = 0;
int i298 = 0;
int i299 = 0;
int i300 = 0;
int i301 = 0;
int i302 = 0;
int i303 = 0;
int i304 = 0;
int i305 = 0;
int i306 = 0;
int i307 = 0;
int i308 = 0;
int i309 = 0;
int i310 = 0;
int i311 = 0;
int i312 = 0;
int i313 = 0;
int i314 = 0;
int i315 = 0;
int i316 = 0;
int i317 = 0;
int i318 = 0;
int i319 = 0;
int i320 = 0;
int i321 = 0;
int i322 = 0;
int i323 = 0;
int i324 = 0;
int i325 = 0;
int i326 = 0;
int i327 = 0;
int i328 = 0;
int i329 = 0;
int i330 = 0;
int i331 = 0;
int i332 = 0;
int i333 = 0;
int i334 = 0;
int i335 = 0;
int i336 = 0;
int i337 = 0;
int i338 = 0;
int i339 = 0;
int i340 = 0;
int i341 = 0;
int i342 = 0;
int i343 = 0;
int i344 = 0;
int i345 = 0;
int i346 = 0;
int i347 = 0;
int i348 = 0;
int i349 = 0;
int i350 = 0;
int i351 = 0;
int i352 = 0;
int i353 = 0;
int i354 = 0;
int i355 = 0;
int i356 = 0;
int i357 = 0;
int i358 = 0;
int i359 = 0;
int i360 = 0;
int i361 = 0;
int i362 = 0;
int i363 = 0;
int i364 = 0;
int i365 = 0;
int i366 = 0;
int i367 = 0;
int i368 = 0;

```


-173-

【図178】

```

// TRUE
printf("Combining Y M N and passing deeper\n");
breaklist_remove_break(blst, bn);
breaklist_remove_break(blst, bn);
b=2;
b = breaklist_return_break_ptr(blst, bn);
break_count = 1;
b = layer_gradient(image, blst, bn, bn+1, thresh_level, 1);
return b;
}
if ((blst[2] == 0) && (blst[2] == 1)) // YNNY
{
printf("Y M N\n");
printf("Trying Y M N = ");
if ((pass_to_deeper_level(image, blst, bn+1, bn+4))
{
printf("Good\n");
breaklist_remove_break(blst, bn+2);
breaklist_remove_break(blst, bn+3);
b = breaklist_return_break_ptr(blst, bn);
break_count = 1;
return b;
}
}
printf("Combining Y M N and passing deeper\n");
breaklist_remove_break(blst, bn);
breaklist_remove_break(blst, bn);
b=2;
b = breaklist_return_break_ptr(blst, bn);
break_count = 1;
b = layer_gradient(image, blst, bn, bn+1, thresh_level, 1);
return b;
}
printf("WARNING: No pattern matched the goodness routines.\n");
return b;
}

// The image is a grey image
// Space allocated for a break list
// This will be initialized
breaklist = blst;
// The first break, last break, // Blank used for break numbers
// The first and last breaks
// first break = 0;
// last break = blst - blst_count - 1;
// If last break is 1;
}

```

```

// TRUE
printf("Combining Y M N and passing deeper\n");
breaklist_remove_break(blst, bn);
breaklist_remove_break(blst, bn);
b=2;
b = breaklist_return_break_ptr(blst, bn);
break_count = 1;
b = layer_gradient(image, blst, bn, bn+1, thresh_level, 1);
return b;
}
if ((blst[2] == 0)
{
printf("Y M N\n");
printf("Trying Y M N = ");
if ((pass_to_deeper_level(image, blst, bn+2, bn+4))
{
printf("Good\n");
breaklist_remove_break(blst, bn+2);
breaklist_remove_break(blst, bn+3);
b = breaklist_return_break_ptr(blst, bn);
break_count = 1;
return b;
}
}
printf("Combining Y M N and passing deeper\n");
breaklist_remove_break(blst, bn);
breaklist_remove_break(blst, bn);
b=2;
b = breaklist_return_break_ptr(blst, bn);
break_count = 1;
b = layer_gradient(image, blst, bn, bn+1, thresh_level, 1);
return b;
}
printf("WARNING: No pattern matched the goodness routines.\n");
return b;
}

// The image is a grey image
// Space allocated for a break list
// This will be initialized
breaklist = blst;
// The first break, last break, // Blank used for break numbers
// The first and last breaks
// first break = 0;
// last break = blst - blst_count - 1;
// If last break is 1;
}

```


[illegible]

【図185】

```

// This routine will insert an entry into the PVP cache (a cyclic cache)
// with indices p1,v1,p2 and output loc,angle.
void L_PVP_cache_insert(
    int p1, // The top of the peak
    int v1, // The start of valley
    int p2, // The end of valley
    int v2, // The top of 2nd peak
    int p3, // The character we're in
    int min_angle, // The final angle
    int min_at) // Where the minimum is at
{
    if (PVP_CACHE_SIZE == 0)
        return;
    int i;
    // Make sure the PVP array doesn't have a one running through it!
    for (i = 0; i < PVP_SIZE; i++)
        if (PVP[i] == 1)
            return;
    if (PVP_SIZE == 0)
        // Insert last element
        PVP_SIZE = 1;
    PVP[0] = 0;
    // Also (
    if (PVP_SIZE == PVP_SIZE) // Not full, wrap last out
        PVP[0] = PVP[0];
    // The next line is basically PVP_SIZE = (PVP_SIZE-1) % PVP_SIZE
    PVP_SIZE = (PVP_SIZE-1) % PVP_SIZE;
    // Also (
    PVP[0] = (PVP_SIZE-1) % PVP_SIZE;
    // Now place the new data at PVP_SIZE
    PVP_SIZE = PVP_SIZE;
    PVP[0] = PVP[0];
    PVP[1] = PVP[1];
    PVP[2] = PVP[2];
    PVP[3] = PVP[3];
    PVP[4] = PVP[4];
    PVP[5] = PVP[5];
    PVP[6] = PVP[6];
    PVP[7] = PVP[7];
    PVP[8] = PVP[8];
    PVP[9] = PVP[9];
    // This routine will check PVP against past results and if we've done
    // it before, it will just grab the results.
    int L_PVP_cache_hit(
        int p1, // The top of the peak
        int v1, // The start of valley
        int p2, // The end of valley
        int p3, // The character we're in
        int min_angle, // The final angle
        int min_at) // Output: Where the minimum is at
    {
        int i;
        if (L_PVP_cache_hit)
            return 0;
        if (PVP_SIZE == 0)
            return 0;
    }
}

```

```

// The PVP cache routines
// This routine will insert an entry into the PVP cache (a cyclic cache)
// with indices p1,v1,p2 and output loc,angle.
void L_PVP_cache_insert(
    int p1, // The top of the peak
    int v1, // The start of valley
    int p2, // The end of valley
    int v2, // The top of 2nd peak
    int p3, // The character we're in
    int min_angle, // The final angle
    int min_at) // Where the minimum is at
{
    if (PVP_CACHE_SIZE == 0)
        return;
    int i;
    // Make sure the PVP array doesn't have a one running through it!
    for (i = 0; i < PVP_SIZE; i++)
        if (PVP[i] == 1)
            return;
    if (PVP_SIZE == 0)
        // Insert last element
        PVP_SIZE = 1;
    PVP[0] = 0;
    // Also (
    if (PVP_SIZE == PVP_SIZE) // Not full, wrap last out
        PVP[0] = PVP[0];
    // The next line is basically PVP_SIZE = (PVP_SIZE-1) % PVP_SIZE
    PVP_SIZE = (PVP_SIZE-1) % PVP_SIZE;
    // Also (
    PVP[0] = (PVP_SIZE-1) % PVP_SIZE;
    // Now place the new data at PVP_SIZE
    PVP_SIZE = PVP_SIZE;
    PVP[0] = PVP[0];
    PVP[1] = PVP[1];
    PVP[2] = PVP[2];
    PVP[3] = PVP[3];
    PVP[4] = PVP[4];
    PVP[5] = PVP[5];
    PVP[6] = PVP[6];
    PVP[7] = PVP[7];
    PVP[8] = PVP[8];
    PVP[9] = PVP[9];
    // This routine will check PVP against past results and if we've done
    // it before, it will just grab the results.
    int L_PVP_cache_hit(
        int p1, // The top of the peak
        int v1, // The start of valley
        int p2, // The end of valley
        int p3, // The character we're in
        int min_angle, // The final angle
        int min_at) // Output: Where the minimum is at
    {
        int i;
        if (L_PVP_cache_hit)
            return 0;
        if (PVP_SIZE == 0)
            return 0;
    }
}

```

```

// The PVP cache routines
// This routine will insert an entry into the PVP cache (a cyclic cache)
// with indices p1,v1,p2 and output loc,angle.
void L_PVP_cache_insert(
    int p1, // The top of the peak
    int v1, // The start of valley
    int p2, // The end of valley
    int v2, // The top of 2nd peak
    int p3, // The character we're in
    int min_angle, // The final angle
    int min_at) // Where the minimum is at
{
    if (PVP_CACHE_SIZE == 0)
        return;
    int i;
    // Make sure the PVP array doesn't have a one running through it!
    for (i = 0; i < PVP_SIZE; i++)
        if (PVP[i] == 1)
            return;
    if (PVP_SIZE == 0)
        // Insert last element
        PVP_SIZE = 1;
    PVP[0] = 0;
    // Also (
    if (PVP_SIZE == PVP_SIZE) // Not full, wrap last out
        PVP[0] = PVP[0];
    // The next line is basically PVP_SIZE = (PVP_SIZE-1) % PVP_SIZE
    PVP_SIZE = (PVP_SIZE-1) % PVP_SIZE;
    // Also (
    PVP[0] = (PVP_SIZE-1) % PVP_SIZE;
    // Now place the new data at PVP_SIZE
    PVP_SIZE = PVP_SIZE;
    PVP[0] = PVP[0];
    PVP[1] = PVP[1];
    PVP[2] = PVP[2];
    PVP[3] = PVP[3];
    PVP[4] = PVP[4];
    PVP[5] = PVP[5];
    PVP[6] = PVP[6];
    PVP[7] = PVP[7];
    PVP[8] = PVP[8];
    PVP[9] = PVP[9];
    // This routine will check PVP against past results and if we've done
    // it before, it will just grab the results.
    int L_PVP_cache_hit(
        int p1, // The top of the peak
        int v1, // The start of valley
        int p2, // The end of valley
        int p3, // The character we're in
        int min_angle, // The final angle
        int min_at) // Output: Where the minimum is at
    {
        int i;
        if (L_PVP_cache_hit)
            return 0;
        if (PVP_SIZE == 0)
            return 0;
    }
}

```


[illegible]

—185—

【図190】

```

c->rec = FAILED_RECOGNITION;
ret_value = IL_GOODOUT_39; left(image,elist,c,ich); // Do the 99's
return ret_value;
}
if (c == ch)
{
c->rec = FAILED_RECOGNITION; // Incorporate next character to left
c = c->prev;
ret_value--;
}
// Check for exits
if (ic == 100) // If we're at the end, exit
return 0;
if (ic == 100) // If we've passed a Y
return 0; // Stop searching
if (IL_GOODOUT_39) // If character is too long
return 0;
}

// This routine works with IL_GOODOUT_39 to check '99's after
// successful recognition. ie. #109) if trim #109) is not
// successful, it will detect and remove some errors caused by the false successful recognition
// and by not checking adjacent characters when can't split at level 5)
int IL_GOODOUT_39_right(
char* image,
struct Character *elist,
struct Character *ch,
struct Character *rch)
{
struct Character *c; // The character to start from
int ret_value; // The leftmost character of the range
int temp_value; // The # of chars removed
ret_value = temp_value = 0;
c = ch;
while(1)
{
c->prev = c->rch; // Get the previous character
temp_value++; // At edge
if (c == rch)
return ret_value;
if (c->rec == FAILED_RECOGNITION) // Return if not a M
return ret_value;
if (ic->level_out_at != max_thresh_level) // Return if not '99'
return ret_value;
if (recognition(image,ch,c))
{
IL_combine_characters(ch,c);
c->rec = FAILED_RECOGNITION;
ret_value = temp_value;
temp_value = 0;
c = ch;
}
}

// This routine will start at character ch and search to left from ch
// grouping characters to see if they're recognized. If so they are
// combined. Note: ich can == 0.
IL_GOODOUT_39_left(
char* image,
struct Character *elist,
struct Character *ch,
struct Character *rch)
{
struct Character *c; // The character to start from
int ret_value; // The leftmost character of the range
int temp_value; // The # of chars removed
ret_value = 0;
c = ch;
while(1)
{
// Start at character we're on
// This loop is 'return'ed out of
if (recognition(image, c, rch)) // c to ch recognized?
{
IL_combine_characters(elist, c, rch); // Combine it into one character (in c)
change image;
}
}
}

```


【図192】

```

!
//////////////////////////////////////////////////
//
void level4(
    c_image line,           // The line
    struct Charlist *clist) // The charlist
{
    int proj_min, proj_max, // Used to compute size of proj
    int i;

    #ifdef VISIBLE
    ma_clr_vin();
    int w;
    Cursor1_X = Cursor1_Y = 10;
    line_plot(Cursor1_X, Cursor1_Y); // Plot input image
    Cursor2_Y = Cursor1_Y + line.sizec + 10; // Final output goes at Cursor2
    Cursor2_X = Cursor1_X;
    Cursor2_Y = Cursor2_Y + line.sizec + 10;
    Cursor3_X = Cursor2_X;
    for (c=0; c<WINDOW_SIZE_X; c++)
        ma_plotxy(c, Cursor2_Y); // Plot horiz line
    Cursor1_Y = Cursor2_Y + 20;
    Cursor2_Y = Cursor3_Y;
    Cursor3_X = Cursor2_X;
    #endif

    if (line.sizec > MAX_CHAR_HEIGHT)
    {
        printf("\n*** ERROR ***\n");
        printf("Line size (%td) is greater than max line size allocated in multilevel\n", line.sizec);
        printf("**** routines (MAX_CHAR_HEIGHT = %d). Increase MAX_CHAR_HEIGHT.\n", MAX_CHAR_HEIGHT);
        exit(3);
    }

    // Allocate Proj and VPP
    i = (NOT_MAX_RANGE/NOT_DEC_RESOLUTION)-1;
    proj_min = -i*(line.sizec);
    proj_max = line.sizec - proj_min;
    Proj = ivector_ranged(proj_min, proj_max);
    VPP = ivector(line.sizec);
    VPP_Zero = 1; // Erase the cache

    /* ..... Program goes here ..... */
    goodout(line, clist);

    #ifdef VISIBLE
    // Print final character list
    li_plot_charlist(line, clist, Cursor2_X, Cursor2_Y);
    #endif
    printf("Displaying final level 4 charlist. Enter 0 to continue? ");
    // print_ma_charlist(clist, clist->first); // ***** DELETE ME *****
    i = auto_input_int();

    // Deallocate
    free_ivector(VPP);
    free_ivector_ranged(Proj, proj_min);
}

```

【図203】

```

/*
 * The routines in this file determine if a character is recognized or not
 * (for segmentation)
 */
.....

#include "ma.h"           /* Include lick's graphics routines */
#include "autoinput.h"

extern int WINDOW_SIZE_Y;

int recognized(a, size_x, size_c) /* Return 0 for not recognized, 1 for
    recognized */
{
    int **a;               /* The character matrix a[row][col] */
    int size_x;            /* The number of rows in a */
    int size_c;            /* The number of columns in a */
    int r,c;               /* The row and column of the image */
    int dummy;             /* A dummy variable */

    #ifdef VISIBLE
    /* This routine will print the char at the screen's bottom left corner */
    for (r=0; r<20; r++) // Loop through the rows */
        for (c=0; c<size_c; c++) // Loop through the cols */
            if ((r >= 0) && (c < size_c) && (c <= 0) && (r < size_x))
                if (a[r][c] != 0) /* If the pixel is on */
                    ma_plotxy(c-10, r + WINDOW_SIZE_Y - size_x - 10);
            else
                ma_unplotxy(c-10, r + WINDOW_SIZE_Y - size_x - 10);
            else
                ma_unplotxy(c-20, r + WINDOW_SIZE_Y - size_x - 10);

    ma_flush(); // Flush the display buffer (to make sure
                // the character gets printed on the screen *)
    #endif

    printf("Character displayed. Enter return value (1 = good, 0 = bad)? ");
    dummy = auto_input_int();
    return dummy;
}

```

[図 193]

```

/* This program contains vector and matrix mallocs */
#include <stdio.h>
#include <stdlib.h>
/* Routine for error during memory allocation */
void allocation_error(
    char *s) /* The error text */
{
    int i;
    printf("Memory allocation error %s\n", s);
    fprintf(stderr, "Error: %s\n", s);
    exit(1);
}

/* ***** VECTOR ***** */
/* Allocate an integer vector of size (0..n-1) */
int *vector(int n) /* Size of vector */
{
    int *v;
    v = (int *)malloc(sizeof(int) * n);
    if (!v) allocation_error("In vector()");
    return v;
}

/* Free an integer vector v of size (0..n-1) */
void free_vector(int *v) /* The vector */
{
    free((int *) v);
}

/* ***** MATRIX ***** */
/* Allocate an float vector of size (0..n-1) */
float *vector(float n) /* Size of vector */
{
    float *v;
    v = (float *)malloc(sizeof(float) * n);
    if (!v) allocation_error("In vector()");
    return v;
}

/* Free an float vector v of size (0..n-1) */
void free_vector(float *v) /* The vector */
{
    free((float *) v);
}

/* ***** MATRIX ***** */
/* Allocate an integer matrix m of size (0..r-1)(0..c-1) */
int **matrix(int r, int c) /* The size of the matrix */
{
    int **m;
    m = (int **)malloc(sizeof(int *) * r);
    if (!m) allocation_error("In matrix()");
    for (int i = 0; i < r; i++) {
        m[i] = (int *)malloc(sizeof(int) * c);
        if (!m[i]) allocation_error("In matrix()");
    }
    return m;
}

/* Free an integer matrix m of size (0..r-1)(0..c-1) */
void free_matrix(int **m) /* The matrix */
{
    int r;
    for (r = 0; r < m[0]; r++) {
        free(m[r]);
    }
    free((int **) m);
}

/* ***** MATRIX ***** */
/* Allocate an char matrix m of size (0..r-1)(0..c-1) */
char **matrix(char r, int c) /* The size of the matrix */
{
    char **m;
    m = (char **)malloc(sizeof(char *) * r);
    if (!m) allocation_error("In matrix()");
    for (int i = 0; i < r; i++) {
        m[i] = (char *)malloc(sizeof(char) * c);
        if (!m[i]) allocation_error("In matrix()");
    }
    return m;
}

/* Free an char matrix m of size (0..r-1)(0..c-1) */
void free_matrix(char **m) /* The matrix */
{
    int r;
    for (r = 0; r < m[0]; r++) {
        free(m[r]);
    }
    free((char **) m);
}

```


【図198】

```

t_turn_right(); // Check area to behind of mouse
if ((t_in_front(line) & 5) == 1) // If something there
{
    t_move_forward(); // Move behind mouse
    continue; // And loop again
}
not_finished = 0; // Else it's a one-pixel image. so exit
line.pixel[t_cursor][t_cursor] = line.pixel[t_cursor][t_cursor] + 1; // Make this spot
}

// Load the info into the Charlist
new_char->stop = stop;
new_char->bot = bot; // (Make it exclusive)
new_char->left = left;
new_char->right = right; // (Make it exclusive)
new_char->type = CHARACTER_PIXEL_OUTLINED;

Charlist_place_new_char(cclist); // Insert the new character

// Mark the character as removed from the list
for (r=0; r<line.size; r++)
for (c=curvel[c]; c < curvel[r]; c++)
    line.pixel[r][c] = line.pixel[r][c] + 1;
}

////////////////////////////////////
//// Main routine to call to add lines to the image //
////////////////////////////////////
void do_outlining() // Returns number of characters added
{
    cimage line; // image of the line
    struct Charlist* cclist; // The character list
    int start_col; // The range to search through (inclusive)
    int stop_col; // (exclusive)

    int x, y; // The row and column we're on

    for (c=stop_col-1; c>=start_col; c--) // Search through column range
    for (r=line.size-1; r>=0; r--) // Scan from bottom to top
    {
        if ((line.pixel[r][c] & 5) == 1) // On pixel that hasn't been outlined
        {
            line.pixel[r][c] = 133;
            print_outlining(line, c);
            extract_outline_char(line, cclist, start_col, stop_col, x, c);
            // Outline the character
        }
    }
}

```

【図208】

```

//
// Filename: test_mult.c
// This file is used to test and develop multilevel2.c
//
#include <stdio.h>
#include <string.h>
#include <image.h>
#include <ymallev.h>
#include <xs.h>
#include <charlist.h>

extern void l4_initialize();
void level4(cimage line, struct Charlist *cclist);

void xm_init()
{
    struct Charlist cclist;
    struct Character *ch;
    cimage image;
    int i, j;

    l4_initialize();

    Charlist_initialize(&cclist, 100, 100); // 10 chars, 100-max height

    char filename1[100];
    char filename2[100];
    printf("Input filename? /d3/xif/");
    scanf("%s", filename1);
    strcpy(filename2, "/d3/xif/");
    strcat(filename2, filename1);
    image_readiff(filename2);
    image_eligspace();

    ch = &(cclist.chr(cclist.size)); // Set new_char to the last character

    // Load the character
    ch->stop = 0;
    ch->bot = image.size;
    ch->left = 0;
    ch->right = image.size;
    for (i=ch->stop; i<ch->bot; i++) {
        ch->curvel[i] = ch->left;
        ch->curve2[i] = ch->right;
    }

    Charlist_place_new_char(&cclist);

    level4(image, &cclist);

    printf("**** End of Program ****\n");
}

```

[1 9 9]

```

//***** This routine will plot a scaled image on the screen *****/
void clear_screen_and_plot_image() // The image
{
    plot_image_scaled();
    int x, y;
    int i, j;
    // Compute scale to fit image on screen (x and y)
    scale = (float) (WINDOW_WIDTH - 10) / (float) image_xsize;
    scale = (float) (WINDOW_HEIGHT - 10) / (float) image_ysize;
    if (scale < scale_x) // Take the least of the two
        scale = scale_x; // Only shrink the image
    if (scale > 1.0)
        scale = 1.0;
    x = (int) ((double) image_xsize * scale);
    y = (int) ((double) image_ysize * scale);
    // Clear the window
    for (i = 0; i < x; i++)
        for (j = 0; j < y; j++)
            plot_image_scaled(i, j);
    // Draw the image
    for (i = 0; i < x; i++)
        for (j = 0; j < y; j++)
            plot_image_scaled(i, j);
    // Draw Box Around Image
    x1 = 0;
    x2 = x;
    y1 = 0;
    y2 = y;
    for (i = 0; i < x; i++)
        for (j = 0; j < y; j++)
            plot_image_scaled(i, j);
    // The input image
    // The starting and stopping col (inclusive)
    // The starting and stopping row (inclusive)
    // The output histogram (allocated previously)
    int i, j;
}

//***** This routine will plot a scaled image on the screen *****/
// This will input a FOR TEXT 0 FOR BLANK!
void make_histogram()
{
    // The input image
    // The starting and stopping col (inclusive)
    // The starting and stopping row (inclusive)
    // The output histogram (allocated previously)
    int i, j;
}

//***** This routine will plot a scaled image on the screen *****/
// This will input a FOR TEXT 0 FOR BLANK!
void make_histogram()
{
    // The input image
    // The starting and stopping col (inclusive)
    // The starting and stopping row (inclusive)
    // The output histogram (allocated previously)
    int i, j;
}

```


[illegible]

—196—

【図206】

```

/* This program will test the charlist */
#include "charlist.h"

void print_charlist(charlist)
struct Charlist *clist;
{
    struct Character *ch;

    printf("Charlist:\n");
    ch = clist->first;

    while(ch != 0) {
        printf("Loc: %d, Chars (%d) = '%c', ch->loc=%d",
            ch->prev == 0 ? 0 : ch->prev,
            ch->next == 0 ? 0 : ch->next,
            ch->char,
            ch->loc);
        printf("\n");
        ch = ch->next;
    }
    printf("End of charlist\n\n");
}

main()
{
    struct Character *ch;
    struct Charlist clist;
    int a[10]; /* Dummy vector for height */
    int i, p;

    Charlist_initialize(&clist, 10, 10);

    for (i=0; i<10; i++) {
        printf("Position of new char %d: ", i);
        scanf("%d", &p);
        ch = &clist->ch[clist->size];
        ch->top = i;
        ch->bot = i-1;
        ch->curvel[1] = ch->curvel[2] = p;
        Charlist_place_new_char(&clist);
        print_charlist(&clist);
    }

    while(1) {
        printf("Replace which char? ");
        scanf("%d", &i);

        ch = clist->first;
        while(ch != 0) {
            if (ch->top == i)
                break;
            ch = ch->next;
        }
        if (ch == 0) {
            printf("Character not found.\n");
        } else {
            printf("New position? ");
        }
    }
}

```

【図209】

```

/* This file prints a character to be recognised on the screen by */
/* calling the routine thieu_recognize() */

void thieu_recognize(size_t, size_t, image)
int size_x; /* The number of rows */
int size_y; /* The number of columns */
char **image; /* The image; bit 1 is the image */
{
    int r, c;

    printf("\nThe character is:\n");
    for (r=0; r<size_y; r++) {
        for (c=0; c<size_x; c++) {
            if (image[r][c] & 1)
                printf(" "); /* On */
            else
                printf("."); /* Off */
        }
        printf("\n");
    }
    printf("***** End of character *****\n\n");
}

```


【図207】

```

// Filename: test_mult.c
// This file is used to test and develop multilevel.c
// =====
//
#include <stdio.h>
#include "images.h"
#include "xmalloc.h"
#include "xs.h"
#include "charlist.h"

extern void l4_initialize();
void level4(image_t line, struct Charlist *clist);

void xs_init()
{
    struct Charlist clist;
    struct Character *ch;
    image image;
    int i,j;

    l4_initialize();
    Charlist_initialize(&clist, 10, 300); // 10 chars, 100-max height
    ch = &(clist.ch[clist.size]); // Set new_char to the last character

    // =====
    // Load the image
    //
    image.fill(1,10,10);
    image.pixel[6][4] = 0;
    image.pixel[7][4] = 0;
    image.pixel[8][4] = 0;
    image.pixel[6][7] = 0;
    image.pixel[7][7] = 0;
    image.pixel[8][7] = 0;
    image.pixel[6][8] = 0;
    image.pixel[7][8] = 0;
    image.pixel[8][8] = 0;
    image.pixel[6][9] = 0;
    image.pixel[7][9] = 0;
    image.pixel[8][9] = 0;
    image.pixel[2][10] = 0;
    image.pixel[3][10] = 0;
    image.pixel[4][10] = 0;
    image.pixel[5][10] = 0;
    image.pixel[6][10] = 0;
    image.pixel[7][10] = 0;
    image.pixel[8][10] = 0;
    image.pixel[2][11] = 0;
    image.pixel[3][11] = 0;
    image.pixel[4][11] = 0;
    image.pixel[5][11] = 0;
    image.pixel[6][11] = 0;
    image.pixel[7][11] = 0;
    image.pixel[8][11] = 0;
    image.pixel[0][12] = 0;
    image.pixel[1][12] = 0;
    image.pixel[2][12] = 0;
    image.pixel[3][12] = 0;
    image.pixel[4][12] = 0;
    image.pixel[5][12] = 0;
    image.pixel[6][12] = 0;
    image.pixel[7][12] = 0;
    image.pixel[8][12] = 0;
    image.pixel[0][13] = 0;
    image.pixel[1][13] = 0;
    image.pixel[2][13] = 0;
}

image.pixel[0][14] = 0;
for (i=1; i<9; i++)
    image.pixel[i][20] = 0;
for (i=0; i<8; i++)
    image.pixel[i][21] = 0;
for (i=1; i<9; i++)
    image.pixel[i][22] = 0;
for (i=27; i<30; i++)
    for (j=0; j<8; j++)
        image.pixel[j][i] = 0;
for (i=28; i<29; i++)
    for (j=1; j<2; j++)
        image.pixel[j][i] = 0;
for (i=29; i<31; i++)
    for (j=0; j<2; j++)
        image.pixel[j][i] = 0;

// Load the character
ch->stop = 0;
ch->bot = 10;
ch->slc = 0;
ch->right = 33;
for (i=ch->stop; i<ch->bot; i++) {
    ch->curve1[i] = ch->left;
    ch->curve2[i] = ch->right;
}

Charlist_place_new_char(&clist);
level4(image, &clist);
printf("**** End of Program ****\n");
}

```

【図214】

```

/* If your routines are in C++ then CPP must be defined in make file */
#ifdef CPP
extern "C" void xs_set_foreground(int x, int y, int b);
extern "C" void xs_set_background(int x, int y, int b);
extern "C" void xs_init();
extern "C" void xs_flush();
extern "C" void xs_getxy(int x, int y);
extern "C" void xs_getsize(int x, int y);
extern "C" void xs_display();
extern "C" void xs_clear_image(int x, int y, int width, int height, char *buf);
extern "C" void xs_clear_win();
extern "C" void xs_draw_box(int x, int y, int width, int height);
extern "C" void xs_string(int x, int y, char *str);
extern "C" void xs_setcolor(int color);
extern "C" void xs_vline(int x, int y, int height);
#endif

```

【図210】

```

/* FILE: util.c
CONCEPT:
This file contains general utility routines.
ut_getbits(width,height,buf);
ut_getbits(nc,n);
ut_getbits(nc,width,height,buf);
Usage:
*/
Just call the routines.
/* .....includes..... */
#include <stdio.h>
#include <stdlib.h>
/* .....global..... */
/* .....externs..... */
/*
Routines: ut_getbits(width,height,buf)
Description: Reverse the bits for display format
Return: Nothing
Author: Rick Lee
*/
ut_getbits(width,height,buf)
/* Reverse the bits for display format
int width;
int height;
char *buf;
{
    unsigned char val,new_val;
    int x,y;
    for(x = 0; x < (width-1)*height; x++)
    {
        val = buf[x];
        new_val = 0;
        for(y = 7; y >= 0; y--)
        {
            if((ut_getbits(nc,y+1))
            {
                new_val = (new_val << 1) | new_val;
            }
            else
            {
                new_val = new_val;
            }
        }
        buf[x] = new_val;
    }
}
/*
Routines: ut_getbits(nc,y,n)
Description: Check if the bit is turned on or not
Return: Nothing
Author: Rick Lee
*/
ut_getbits(nc,y,n)
{
    unsigned char *p;
    int bit;
    /* The character to give bit is from */
    /* The beginning of the bit number */
    /* Number of bits to get */
    return(x >> ((p-1-n) & -(1 << n));
}
/*
Routines: ut_getbits(nc,width,height,buf)
Description: Display the bitmap using ascii characters
Return: Nothing
Author: Rick Lee
*/
ut_getbits(nc,width,height,buf)
/* Display the bitmap using ascii characters
int width;
int height;
char *buf;
{
    short bit_code = 0x0010;
    short bit_result = 0;
    short x,y,s_bits;
    for(x = 0; x < height; x++)
    {
        for(y = 0; y < width; y++)
        {
            bits = buf[(x*y)];
            for(s = 0; s < 8; s++)
            {
                bit_result = bits & bit_code;
                if(bit_result != 0)
                {
                    printf("\n");
                }
                bit_code = bit_code >> 1;
            }
            bit_code = 0x0010;
        }
        printf("\n");
    }
}

```


[図212]

```

/* Routine: re_display()
Description: Redisplay the "stuff" that was on the window
Returns: Nothing
Author: Rick Lee
*/
re_display()
{
    /* Routine: re_display(canvas), xWindow(canvas), 0.0,0.0,TRUE;
    */
    /* Routine: re_draw_box(x,y,width,height)
    Description: Draw a box on screen
    Returns: Nothing
    Author: Rick Lee
    */
    re_draw_box(x,y,width,height);
    int x;
    int y;
    int width;
    int height;
    {
        XDrawImageString(canvas,x,y,width,height);
    }
}

/* Routine: re_string
Description: Draw a string on the screen
Returns: Nothing
Author: Rick Lee
*/
re_string(x,y, str)
int x;
int y;
char *str;
{
    XDrawImageString(canvas,x,y,width,height);
}

/* Routine: re_setcolor
Description: Set the current color
Returns: Nothing
Author: Rick Lee
*/
re_setcolor(color)
int color;
{
    (cv.foreground = color);
    XSetForeground(canvas,x,y,width,height);
}

```

```

/* Routine: re_display()
Description: Redisplay the "stuff" that was on the window
Returns: Nothing
Author: Rick Lee
*/
re_display()
{
    /* Routine: re_draw_box(x,y,width,height,buf)
    Description: Draw the given bitmap
    Returns: Nothing
    Author: Rick Lee
    */
    re_draw_box(x,y,width,height,buf);
    /* Create the bitmap area */
    pix = XCreatePixmap(canvas,xWindow(canvas),width,height,depth);
    /* Reverse the bit ordering */
    re_reverse(width,height,buf);
    /* First create bitmap */
    bitmap = XCreateBitmapFromData(canvas,xWindow(canvas),width,height,buf,width,height);
    /* Create the pixmap area */
    pix = XCreatePixmap(canvas,xWindow(canvas),width,height,depth);
    /* Copy the data to bitmap */
    XCopyArea(canvas,xWindow(canvas),pix,width,height,0,0,0,0,0);
    /* Now copy the pixmap to drawing area */
    XCopyArea(canvas,xWindow(canvas),pix,xWindow(canvas),line_gc,0,0,width,height,x,y);
    /* Now free the pixmap */
    XFreePixmap(canvas,bitmap);
    /* Now free the bitmap */
    XFreeBitmap(canvas,bitmap);
    /* Now finish the queue to X-server */
    XFlush(canvas);
}

/* Routine: re_display()

```

【図213】

```

/*
Routine: xs_win_dia
Description: Get the size of the window
Returns: Nothing
Author: Rich Lee
*/
xs_wdia(wid, height)
int *wid;
int *height;
{
    Arg args[2];
    Dimension win_width, win_height;

    XSetArg(args[0], XcWidth, win_width);
    XSetArg(args[1], XcHeight, win_height);
    XGetValues(canvas, args, 2);

    *wid = win_width;
    *height = win_height;
}

/*
xs_set_map[w]
Widget w;
{
    int ncolors;
    XColor colors[256];
    Colormap my_map;
    int count = 1;

    Display *d;
    int scr;
    Colormap def_colormap;
    XColor Color;
    int x;
    int the_color = 1000;

    d = XOpenDisplay(w);
    scr = DefaultScreen(d);
    ncolors = DisplayCells(d, scr);
    def_colormap = DefaultColormap(d, scr);
    for(x = 0; x < ncolors; x++)
    {
        colors[x].pixel = x;
        colors[x].flags = DoRed | DoGreen | DoBlue;

        colors[x].red = the_color;
        colors[x].green = the_color;
        colors[x].blue = the_color;
        the_color += 100;
    }
    XQueryColors(d, def_colormap, colors, ncolors);

    my_map = XCreateColormap(d, DefaultRootWindow(d), DefaultVisual(d, scr),
        AllocAll);
    XStoreColors(d, my_map, colors, ncolors);
}
*/

```

【図228】

```

c1 = get2bytes(f, fp);
c2 = get2bytes(f, fp);
return c1 - c2-65536;
}

// Plot image using Flash Graphics at (xlib, ylib) with color
void image::plot(int x, int y)
{
    int px, py;
    char color;

    for (py=0; py<sizeR; py++)
        for (px=0; px<sizeC; px++)
        {
            color = pixel[py][px];
            if (color & 1)
                XDrawPixel(x-px, y-py); // Draw it
        }
}

// Plot image with a box around it (with color 1)
void image::plotbox(int x, int y)
{
    int px, py;

    for (px=0; px<sizeC; px++) { // Draw horizontal lines
        if ((x-px-y)>2 && 0)
            XDrawPixel(x-px, y);
        if ((x-px-y)<-2 && 0)
            XDrawPixel(x-px, y+sizeR-1);
    }
    for (py=0; py<sizeR; py++) {
        if ((x-y-py)>2 && 0)
            XDrawPixel(x, y-py);
        if ((x-y-py)<-2 && 0)
            XDrawPixel(x+sizeC-1, y-py);
    }
    plot(x-1, y-1); // Plot the image
}

```

【図240】

```

/* This file prints a character to be recognized on the screen by */
/* calling the routine thic_recognise() */

char thic_recognise(size_r, size_c, image)
int size_r; // The number of rows
int size_c; // The number of columns
char **image; // The image: 2it 1 is the image
{
    int c;
    char input_string[100];

    printf("\nThe character is:\n");
    for (r=0; r<size_r; r++)
    {
        for (c=0; c<size_c; c++)
        {
            if (image[r][c] & 1)
                printf(" "); // On
            else
                printf("."); // Off
        }
        printf("\n");
    }
    printf("***** End of character *****\n");
    printf("Enter Character? ");
    scanf("%s", input_string);
    return input_string[strlen(input_string)-1];
}

```


【図216】

```

// filename: zerovpp.c
// This routine contains the lines to find blocks that have non-zero vpp
// (level 1 segmentation)
//
// =====
//
#include "charlist.h"
#include "images.h"
// =====
// Scan a line quickly (every 1/4 pixel) looking for an 'on' pixel
int line_scan_quick( // Returns 1 for pixel found, 0 otherwise
    images_line,      // The line to process
    int c)             // The column to scan
{
    int r;
    for (r=0; r<line.size; r+=4)
        if (line.pixel(r)(c) & 1)
            return 1;
    return 0;
}

// =====
// Scan a line completely looking for an 'on' pixel
int line_scan_complete( // Returns 1 for pixel found, 0 for otherwise
    images_line,         // The line
    int c)               // The column to scan
{
    int r;
    for (r=0; r<line.size; r++)
        if (line.pixel(r)(c) & 1)
            return 1;
    return 0;
}

// =====
// The main routine. Call this to make the list
void extract_zero_vpp_list(
    images_line, // The image of the line to segment
    struct NonZeroVppList *list) // A pointer to the list
{
    int c;
    int left, right;
    for (c=0; c<line.size; c++) // Scan the image
    {
        if (line_scan_quick(line, c)) // If found something
        {
            left = right = c; // Look to the left for the start
            while (--c >= 0)
                if (line_scan_complete(line, c) &= 0)
                    break;
            left = c+1;
            c = right; // Keep looking to the right
            while (++c < line.size)
                if (line_scan_complete(line, c) &= 0)
                    break;
            right = c;
            zerovpp_add(list, left, right); // Add in the new break
        }
    }
}

```

【図245】

```

/* If your routines are in C++ then CPP must be defined in make file */
#ifdef CPP
extern "C" void xa_init();
extern "C" void xa_flush();
extern "C" void xa_plotxy(int x, int y);
extern "C" void xa_unplotxy(int x, int y);
extern "C" void xa_redisplay();
extern "C" void xa_display_image(int x, int y, int width, int height, char *buf);
extern "C" void xa_clr_screen();
extern "C" void xa_draw_box(int x, int y, int width, int height);
extern "C" void xa_string(int x, int y, char *str);
extern "C" void xa_setcolor(int color);
extern "C" void xa_vin_dim(int *width, int *height);
#endif

```


[図219]

```

/* filename: autoinput.h
 * This file contains the code to input from a filename
 * .....
 */
#include <stdio.h>
#include "autoinput.h"
/* Global variables */
int auto_eof_detected;
FILE *auto_fp;
int save_to_file;
int save_count;
char filename[100];
/* .....
void auto_initialize()
{
    int not_exit = 1;
    auto_eof_detected = 0;
    while(not_exit)
    {
        printf("Input filename for autoinput? ");
        scanf("%s", filename);
        auto_fp = fopen(filename, "r");
        if (auto_fp == NULL) /* lead open failed */
        {
            auto_fp = fopen(filename, "w");
            if (auto_fp == NULL) /* write open failed */
            {
                printf("cannot open file %s\n", filename);
                printf("Please check the filename and file\n");
                printf("number 1 at the next prompt.\n");
                save_to_file = 1;
                auto_eof_detected = 1;
                not_exit = 0;
            }
            else {
                printf("opening existing file %s\n", filename);
                not_exit = 0;
            }
        }
        printf("Do you wish to save the auto-input into this file? (1 = yes) ");
        save_to_file = auto_input_int();
        printf("Warning: Change the file '1' to a '0' to turn off appending to the file.\n");
        save_count = 0;
    }
}

int auto_input_int()
{
    int i;
    if (auto_eof_detected)
    {
        printf("\nWarning: Input (-9 = newline, -9999 = exit) ? ");

```

```

scanf("%d", &i);
if (i == -9999)
{
    fclose(auto_fp);
    exit(0);
}
if (save_to_file == 1)
{
    if (i == -9) /* Print new lines */
    {
        printf(auto_fp, "\n"); /* Print the new lines */
        fflush(auto_fp);
        save_count = 0;
        return auto_input_int(); /* And input again */
    }
    printf(auto_fp, "%d ", i);
    fflush(auto_fp);
    if (++save_count > 10)
    {
        save_count = 0;
        printf(auto_fp, "\n");
    }
    return i;
}
if (fclose(auto_fp, "w", &i) == EOF)
{
    printf("\n----- END OF FILE DETECTED -----");
    if (save_to_file == 1)
    {
        fclose(auto_fp);
        auto_fp = fopen(filename, "a");
        if (auto_fp == NULL)
        {
            printf("----- WARNING: CANNOT APPEND TO FILE, IS 'a' NOT\n", filename);
            save_to_file = 0;
        }
        printf(auto_fp, "\n");
        auto_eof_detected = 1;
        return auto_input_int();
    }
    printf("\n\n");
    return i;
}

```

```

1  * filename: charlist.c
2  * This file contains CharList and HandleCharList utility routines
3  *
4  *
5  *
6  *
7  *
8  *
9  *
10 *
11 *
12 *
13 *
14 *
15 *
16 *
17 *
18 *
19 *
20 *
21 *
22 *
23 *
24 *
25 *
26 *
27 *
28 *
29 *
30 *
31 *
32 *
33 *
34 *
35 *
36 *
37 *
38 *
39 *
40 *
41 *
42 *
43 *
44 *
45 *
46 *
47 *
48 *
49 *
50 *
51 *
52 *
53 *
54 *
55 *
56 *
57 *
58 *
59 *
60 *
61 *
62 *
63 *
64 *
65 *
66 *
67 *
68 *
69 *
70 *
71 *
72 *
73 *
74 *
75 *
76 *
77 *
78 *
79 *
80 *
81 *
82 *
83 *
84 *
85 *
86 *
87 *
88 *
89 *
90 *
91 *
92 *
93 *
94 *
95 *
96 *
97 *
98 *
99 *
100 *
101 *
102 *
103 *
104 *
105 *
106 *
107 *
108 *
109 *
110 *
111 *
112 *
113 *
114 *
115 *
116 *
117 *
118 *
119 *
120 *
121 *
122 *
123 *
124 *
125 *
126 *
127 *
128 *
129 *
130 *
131 *
132 *
133 *
134 *
135 *
136 *
137 *
138 *
139 *
140 *
141 *
142 *
143 *
144 *
145 *
146 *
147 *
148 *
149 *
150 *
151 *
152 *
153 *
154 *
155 *
156 *
157 *
158 *
159 *
160 *
161 *
162 *
163 *
164 *
165 *
166 *
167 *
168 *
169 *
170 *
171 *
172 *
173 *
174 *
175 *
176 *
177 *
178 *
179 *
180 *
181 *
182 *
183 *
184 *
185 *
186 *
187 *
188 *
189 *
190 *
191 *
192 *
193 *
194 *
195 *
196 *
197 *
198 *
199 *
200 *
201 *
202 *
203 *
204 *
205 *
206 *
207 *
208 *
209 *
210 *
211 *
212 *
213 *
214 *
215 *
216 *
217 *
218 *
219 *
220 *
221 *
222 *
223 *
224 *
225 *
226 *
227 *
228 *
229 *
230 *
231 *
232 *
233 *
234 *
235 *
236 *
237 *
238 *
239 *
240 *
241 *
242 *
243 *
244 *
245 *
246 *
247 *
248 *
249 *
250 *
251 *
252 *
253 *
254 *
255 *
256 *
257 *
258 *
259 *
260 *
261 *
262 *
263 *
264 *
265 *
266 *
267 *
268 *
269 *
270 *
271 *
272 *
273 *
274 *
275 *
276 *
277 *
278 *
279 *
280 *
281 *
282 *
283 *
284 *
285 *
286 *
287 *
288 *
289 *
290 *
291 *
292 *
293 *
294 *
295 *
296 *
297 *
298 *
299 *
300 *
301 *
302 *
303 *
304 *
305 *
306 *
307 *
308 *
309 *
310 *
311 *
312 *
313 *
314 *
315 *
316 *
317 *
318 *
319 *
320 *
321 *
322 *
323 *
324 *
325 *
326 *
327 *
328 *
329 *
330 *
331 *
332 *
333 *
334 *
335 *
336 *
337 *
338 *
339 *
340 *
341 *
342 *
343 *
344 *
345 *
346 *
347 *
348 *
349 *
350 *
351 *
352 *
353 *
354 *
355 *
356 *
357 *
358 *
359 *
360 *
361 *
362 *
363 *
364 *
365 *
366 *
367 *
368 *
369 *
370 *
371 *
372 *
373 *
374 *
375 *
376 *
377 *
378 *
379 *
380 *
381 *
382 *
383 *
384 *
385 *
386 *
387 *
388 *
389 *
390 *
391 *
392 *
393 *
394 *
395 *
396 *
397 *
398 *
399 *
400 *
401 *
402 *
403 *
404 *
405 *
406 *
407 *
408 *
409 *
410 *
411 *
412 *
413 *
414 *
415 *
416 *
417 *
418 *
419 *
420 *
421 *
422 *
423 *
424 *
425 *
426 *
427 *
428 *
429 *
430 *
431 *
432 *
433 *
434 *
435 *
436 *
437 *
438 *
439 *
440 *
441 *
442 *
443 *
444 *
445 *
446 *
447 *
448 *
449 *
450 *
451 *
452 *
453 *
454 *
455 *
456 *
457 *
458 *
459 *
460 *
461 *
462 *
463 *
464 *
465 *
466 *
467 *
468 *
469 *
470 *
471 *
472 *
473 *
474 *
475 *
476 *
477 *
478 *
479 *
480 *
481 *
482 *
483 *
484 *
485 *
486 *
487 *
488 *
489 *
490 *
491 *
492 *
493 *
494 *
495 *
496 *
497 *
498 *
499 *
500 *
501 *
502 *
503 *
504 *
505 *
506 *
507 *
508 *
509 *
510 *
511 *
512 *
513 *
514 *
515 *
516 *
517 *
518 *
519 *
520 *
521 *
522 *
523 *
524 *
525 *
526 *
527 *
528 *
529 *
530 *
531 *
532 *
533 *
534 *
535 *
536 *
537 *
538 *
539 *
540 *
541 *
542 *
543 *
544 *
545 *
546 *
547 *
548 *
549 *
550 *
551 *
552 *
553 *
554 *
555 *
556 *
557 *
558 *
559 *
560 *
561 *
562 *
563 *
564 *
565 *
566 *
567 *
568 *
569 *
570 *
571 *
572 *
573 *
574 *
575 *
576 *
577 *
578 *
579 *
580 *
581 *
582 *
583 *
584 *
585 *
586 *
587 *
588 *
589 *
590 *
591 *
592 *
593 *
594 *
595 *
596 *
597 *
598 *
599 *
600 *
601 *
602 *
603 *
604 *
605 *
606 *
607 *
608 *
609 *
610 *
611 *
612 *
613 *
614 *
615 *
616 *
617 *
618 *
619 *
620 *
621 *
622 *
623 *
624 *
625 *
626 *
627 *
628 *
629 *
630 *
631 *
632 *
633 *
634 *
635 *
636 *
637 *
638 *
639 *
640 *
641 *
642 *
643 *
644 *
645 *
646 *
647 *
648 *
649 *
650 *
651 *
652 *
653 *
654 *
655 *
656 *
657 *
658 *
659 *
660 *
661 *
662 *
663 *
664 *
665 *
666 *
667 *
668 *
669 *
670 *
671 *
672 *
673 *
674 *
675 *
676 *
677 *
678 *
679 *
680 *
681 *
682 *
683 *
684 *
685 *
686 *
687 *
688 *
689 *
690 *
691 *
692 *
693 *
694 *
695 *
696 *
697 *
698 *
699 *
700 *
701 *
702 *
703 *
704 *
705 *
706 *
707 *
708 *
709 *
710 *
711 *
712 *
713 *
714 *
715 *
716 *
717 *
718 *
719 *
720 *
721 *
722 *
723 *
724 *
725 *
726 *
727 *
728 *
729 *
730 *
731 *
732 *
733 *
734 *
735 *
736 *
737 *
738 *
739 *
740 *
741 *
742 *
743 *
744 *
745 *
746 *
747 *
748 *
749 *
750 *
751 *
752 *
753 *
754 *
755 *
756 *
757 *
758 *
759 *
760 *
761 *
762 *
763 *
764 *
765 *
766 *
767 *
768 *
769 *
770 *
771 *
772 *
773 *
774 *
775 *
776 *
777 *
778 *
779 *
780 *
781 *
782 *
783 *
784 *
785 *
786 *
787 *
788 *
789 *
790 *
791 *
792 *
793 *
794 *
795 *
796 *
797 *
798 *
799 *
800 *
801 *
802 *
803 *
804 *
805 *
806 *
807 *
808 *
809 *
810 *
811 *
812 *
813 *
814 *
815 *
816 *
817 *
818 *
819 *
820 *
821 *
822 *
823 *
824 *
825 *
826 *
827 *
828 *
829 *
830 *
831 *
832 *
833 *
834 *
835 *
8
```

【図222】

```

/* This routine will insert the character beyond the charlist (charlist).
 * loc is calculated for both character and the new character; however
 * loc is NOT USED TO PLACE THE CHARACTER. It will
 * place the new character after character. After finished, it is an
 * extremely good idea to reset the breaklist
 */
void Charlist_Insert_After(charlist, char)
{
    struct Charlist *ch; /* Place the new character after this one */
    struct Charlist *newchar; /* Pointer to the new character adding */
    newchar = (struct Charlist *) malloc(sizeof(struct Charlist)); /* Set newchar */
    /* Calculate the position of the new character */
    Charlist_Calculate_Loc(newchar);
    /* Check for full list */
    if (charlist->next == NULL) /* Increment the size of the list */
    {
        print("Error in Charlist_place_new_char(): Character list is FULL!\n");
        exit(1);
    }
    newchar->prev = charlist;
    newchar->next = charlist->next;
    charlist->next = newchar;
    if (newchar->next != 0)
        newchar->next->prev = newchar;
}

/* This routine will remove the character from the list.
 * Note that only the character is removed, not the pointers on this too much.
 * Also, prev and next pointers of char are NOT altered!
 */
void Charlist_Remove(charlist, ch)
{
    struct Charlist *ch;
    struct Charlist *ch;
    if (ch->prev == 0) /* If list on list */
        charlist->first = charlist->next;
    else
        charlist->prev->next = charlist->next;
    if (ch->next != 0)
        ch->next->prev = ch->prev;
}

/* This routine will recalculate the position of a character */
void Charlist_Rcalculate_Loc(ch)
{
    struct Charlist *ch;
    int i;
    /* Calculate the position of the new character */
    loc = 0;
    for (i = 0; i < charlist->next; i++) /* Position = average position.. */
        loc = (charlist->first + charlist->next) / 2; /* .. of the curve */
    loc = loc * 4 / (charlist->next - charlist->first); /* Set it */
    ch->loc = loc;
}

```

```

/* This routine will place a character (that is not in the charlist)
 * void Charlist_Insert_After(charlist, char)
 * struct Charlist *ch; /* The character */
 * int loc; /* Used for searching for location */
 * struct Charlist *last; /* The character before p */
 * /* Is the list empty? */
 * if (charlist->next == 0)
 * {
 *     ch->next = 0;
 *     charlist->first = ch;
 *     return;
 * }
 * /* Possible optimization, don't start from start of list...
 * loc = ch->loc; /* Retrieve position */
 * /* Is it at the beginning of the list? */
 * p = charlist->first;
 * /* Load p with the first item on the list */
 * /* If it's the first on the list */
 * {
 *     p->prev = ch;
 *     ch->next = charlist->first;
 *     charlist->first = ch;
 *     /* Set the pointers */
 *     /* Set the loc on the list to this */
 *     return;
 * }
 * /* Until find record past where to place */
 * while (p->loc <= loc)
 * {
 *     last = p;
 *     /* Remember last position */
 *     /* Get the next character */
 *     if (p->next == 0)
 *         break;
 *     ch->prev = last;
 *     last = p;
 *     /* Insert new position after last */
 *     if (p->next != 0)
 *     {
 *         p->next->prev = ch;
 *         p->next = ch;
 *     }
 * }
 * /* Allocate memory for list */
 * size = size; /* Compute memory allocated, etc */
 * newlist = (struct Charlist *) malloc(sizeof(struct Charlist)); /* Allocate memory */
 * newlist->first = charlist->first; /* Set new allocated */
 * newlist->next = 0; /* Set to zero elements */
 * newlist->prev = 0;
 * int size; /* The number of entries (is allocated) */

```


【図 2 2 6】

```

for (j=C2; j<=C2; j++)
    temp_pixel[(i-R1)*(j-C2) + pixel(i,j)];
// Copy temp image to current image
for (i=0; i<size; i++)
    for (j=0; j<size; j++)
        pixel[i][j] = temp_pixel[i][j];
}

// Read in part of a TIFF file
void climage::readoff(char *filename)
{
    FILE *tagfile; // Input file
    long int tagfile_position; // Output file position (in bytes)
    long int width; // Width of picture
    long int length; // Length of picture
    long int black; // Value of black pixel
    long int hps; // Bytes per strip
    long int raster_offset; // Start of image data
    long int raster; // Tag values
    long int i,j;

    if ((tagfile = fopen(filename, "rb")) == NULL)
    {
        cout << "Error: TIFF file not in Intel format!\n";
        exit(1);
    }
    tagfile_position = 0;

    // Read in header and make sure it's Intel
    if (fgetc(tagfile) != 'I')
        cout << "... Error: TIFF file not in Intel format!\n";
        exit(1);

    // Read in Magic number
    i = getbytes(tagfile, tagfile_position);
    if (i != 0x0001)
        cout << "BAD Magic Number for TIFF file!\n";
        exit(1);

    // Find first file directory
    i = getbytes(tagfile, tagfile_position);
    while (tagfile_position < j)
    {
        i = getbytes(tagfile, tagfile_position);
        // Tags in Directory
        numtag = getbytes(tagfile, tagfile_position);
        while (numtag > 0)
        {
            tag = getbytes(tagfile, tagfile_position);
            type = getbytes(tagfile, tagfile_position);
            len = getbytes(tagfile, tagfile_position);
            value = getbytes(tagfile, tagfile_position);
            if ((len & 1) && (tag & 32768) == 0)
                cout << "Error: Tag length is 1!\n";
                exit(1);
        }
    }
}

```

```

// MAIN the new image (size size adjusted)
for (int i=0; i<size; i++)
    for (int j=0; j<size; j++)
        pixel[i][j] = value;

// Copy a section of a climage into another. Image = reference image
// Range is [R1..R2] [C1..C2]
void climage::clip(climage *img, int R1, int R2, int C1, int C2)
{
    int R,C,C1;

    if (R2 < R1)
        // Swap R1 & R2
        int t = R1; R1 = R2; R2 = t;
    if (C2 < C1)
        // Swap C1 & C2
        int t = C1; C1 = C2; C2 = t;
    if (R1 < 0) R1 = 0;
    if (R2 > size) R2 = size;
    if (C1 < 0) C1 = 0;
    if (C2 > size) C2 = size;
    if (R1 > R2) R1 = R2;
    if (R2 > size) R2 = size;
    if (C1 > C2) C1 = C2;
    if (C2 > size) C2 = size;

    allocate(R2-R1, C2-C1);

    for (int i=R1; i<R2; i++)
        for (int j=C1; j<C2; j++)
            pixel[(i-R1)*(C2-C1) + j-C1] =
                img->pixel[(i-R1)*(size-size) + j-C1];
}

// Clip space, reduce the size of an image by removing the blank
// space around it. This adjusts the image size.
void climage::clip_image()
{
    int R1, R2, C1, C2; // The starting and stepping rows of new image
    int i,j;
    struct climage temp; // Temporary image

    R1 = size;
    R2 = 0;
    C1 = size;
    C2 = 0;

    for (i=0; i<size; i++)
        for (j=0; j<size; j++)
        {
            if (i < R1) R1 = i;
            if (i > R2) R2 = i;
            if (j < C1) C1 = j;
            if (j > C2) C2 = j;
        }

    if (R1 > R2) R1 = R2;
    if (R2 > size) R2 = size;
    if (C1 > C2) C1 = C2;
    if (C2 > size) C2 = size;

    temp.allocate(R2-R1+1, C2-C1+1); // New image is this size
    for (i=R1; i<R2; i++)
        for (j=C1; j<C2; j++)
            temp.pixel[i-R1][j-C1] = pixel[i][j];
}

```

【図227】

```

while(tagfile_position + raster_offset)
{
    l = getbytes(tagfile_tagfile_position);
}
//----- READ IN FILE -----
allocate(dst, length, (int) width);
int v;
for (i = 0; i < length; i++)
{
    for (j = 0; j < width; j++)
    {
        if ((i+j) == 0)
        {
            v = getbytes(file);
            tagfile_position++;
            if (v == EOF)
            {
                error = "% Error: Premature End of File found in tiff file\n";
                exit(1);
            }
            if ((v & 128) == black*128)
            {
                pixel[i][j] = 0;
            }
            else
            {
                pixel[i][j] = 1;
            }
            v = v & 1;
        }
    }
}
fclose(tagfile);
//----- READ IN A FILE. File position is updated
// error = 1 to error on EOF, 0 not
int cl = getbytes(file);
if (cl == EOF)
{
    error = "% Error: Premature End of File (in getbytes)\n";
    exit(1);
}
return cl;
}
//----- READ IN AN INTEGER FROM A FILE. (LSB, MSB)
long int cl = getbytes(file);
return cl;
}
//----- READ IN A LONG INTEGER
long int cl = getbytes(file);
return cl;
}
//----- READ IN A LONG INTEGER
long int cl = getbytes(file);
return cl;
}

```

【図229】

```

// Filename: cimages.h
// This is the header file for cimages.C
//
//
// =====
// #define MDX_DEBUG 1
#include <stdio.h>
#include <stdlib.h>
#include <math.h>

class cimage; // Forward reference

//-----
//--- CLASS: cimage
//-----
class cimage
{
public:
    int      sizeR; // Number of rows (first index of array)
    int      sizeC; // Number of cols. (second index of array)
    int      allocated_sizeR; // Amount allocated
    int      allocated_sizeC; // Amount allocated
    int      dpi; // Dots per inch of the image (0 if unknown)
    unsigned char **pixels; // The binary image (2d array)

    cimage(); // Initialize
    ~cimage(); // Destroy
    cimage(int r, int c); // Initialize with a size
    cimage(cimage& temp);

    void reallocate(int R, int C); // Will allocate memory for an image
    void fill(char value, int R, int C); // Fill image with a constant value
    void clip(cimage& image, int R1, int R2, int C1, int C2);
    void clipmask(); // Clips zero space surrounding image
    void readtiff(char *filename); // Read in a tiff image
    void plot(int x, int y); // Plot binary image
    void plotbox(int x, int y); // Plot with a box around it

private:
    void allocate(int r, int c); // Allocate memory
    void deallocate(); // Deallocate memory
    int getbyte(FILE *f, long int& file_position, int &error);
    long int getbytes(FILE *f, long int& fpi);
    long int getbytes(FILE *f, long int& fpi);
};

```


【図232】

```

// Copy large into matrix p and call cshl_recompute
{ if(row_bot == 1) at(row_top + 1);
  size_x = row_bot - row_top;
  size_y = col_right - col_left;
  p = cshl_mat(size_x, size_y);
  for (row = row_top; row < row_bot; row++)
  { for (col = col_left; col < col_right; col++)
    { p[row - row_top][col - col_left] = cshl_mat[row - row_top][col - col_left];
    }
  }
  free_matrix(p, size_x);
}

//def visible
//def PAGE_TO_PRINT
  cursor_y += 10; print_vpp_jlist_char((("line_to_process"), cellist,
  what_char, cursor_x, cursor_y));
  while (cursor_y < 10 * print_vpp_jlist((("line_to_process"), cellist, cursor_x, cursor_y));
  {
    readf();
    x = flush();
    printf("Displaying final segmentation. Near 9 to continue? ");
    int done;
    done = auto_input_int();
    zeropp_finalize(maternalist); // ... DELETE ...
    zeropp_finalize(fourcushillist); // ... DELETE ...
  }
}

//def visible
//def PAGE_TO_PRINT
  cursor_y += 10; print_vpp_jlist_char((("line_to_process"), cellist,
  what_char, cursor_x, cursor_y));
  while (cursor_y < 10 * print_vpp_jlist((("line_to_process"), cellist, cursor_x, cursor_y));
  {
    readf();
    x = flush();
    printf("Displaying final segmentation. Near 9 to continue? ");
    int done;
    done = auto_input_int();
    zeropp_finalize(maternalist); // ... DELETE ...
    zeropp_finalize(fourcushillist); // ... DELETE ...
  }
}

```

【図233】

```

/* This program contains common vector and matrix macros */
#include <stdio.h>
#include <stdlib.h>
/* Routine for error: out of memory allocation */
void allocation_error()
{
    printf("Error: out of memory allocation\n");
    exit(1);
}

/* Free an integer ranged vector v of size [n..n-1] */
void free_int_vector(int n)
{
    free((int *) v);
}

/* Allocate an integer matrix m of size [0..r-1][0..c-1] */
int **matrix(int r, int c)
{
    int **m;
    m = (int **) malloc((unsigned) r * sizeof(int *));
    if (!m) allocation_error("in matrix()");
    for (int i = 0; i < r; i++)
        m[i] = (int *) malloc((unsigned) c * sizeof(int));
    if (!m[i]) allocation_error("in matrix()");
    return m;
}

/* Free an integer matrix m of size [0..r-1][0..c-1] */
void free_matrix(int r, int c)
{
    free((int **) m);
}

/* Allocate an float vector of size [0..n-1] */
float *vector(int n)
{
    float *v;
    v = (float *) malloc((unsigned) n * sizeof(float));
    if (!v) allocation_error("in vector()");
    return v;
}

/* Free an float vector v of size [0..n-1] */
void free_vector(int n)
{
    free((float *) v);
}

/* Allocate an float matrix m of size [0..r-1][0..c-1] */
float **matrix(int r, int c)
{
    float **m;
    m = (float **) malloc((unsigned) r * sizeof(float *));
    if (!m) allocation_error("in matrix()");
    for (int i = 0; i < r; i++)
        m[i] = (float *) malloc((unsigned) c * sizeof(float));
    if (!m[i]) allocation_error("in matrix()");
    return m;
}

/* Free an float matrix m of size [0..r-1][0..c-1] */
void free_matrix(int r, int c)
{
    free((float **) m);
}

/* Allocate an char matrix m of size [0..r-1][0..c-1] */
char **matrix(int r, int c)
{
    char **m;
    m = (char **) malloc((unsigned) r * sizeof(char *));
    if (!m) allocation_error("in matrix()");
    for (int i = 0; i < r; i++)
        m[i] = (char *) malloc((unsigned) c * sizeof(char));
    if (!m[i]) allocation_error("in matrix()");
    return m;
}

/* Free an char matrix m of size [0..r-1][0..c-1] */
void free_matrix(int r, int c)
{
    free((char **) m);
}

```

[illegible]

【図239】

```

inc border_ignore;
int i;

charlist_initialize(&clist, MAX_CHARS_PER_LINE, MAX_LINE_HEIGHT);
Process_line.fill(0,255); // Allocate the memory for the lines

xs_win_dim(WINDOW_SIZE_X, WINDOW_SIZE_Y); // Get the window size

printf("Input tiff filename? /d:/tiff/");
strcpy(filename1, "/d:/tiff/test1.tiff");
scanf("%s", filename1);
strcpy(filename, "/d:/tiff/");
strcat(filename, filename1);

auto_initialize(); // Initialize Auto-input routines

printf("Enter courtcut average width? ");
COURCUT_AVO_WIDTH = auto_input_int();

printf("Enter courtcut delta? ");
COURCUT_DELTA = auto_input_int();

page.dpi = 100;

border_ignore = 1;
// cout << "Input border width to ignore (in pixels)? ";
// cin >> border_ignore;

printf("Reading in File... ");
page.readtiff(filename);
printf("Done.\n");

// segm_init(page); // Initialize page printing routines

// border_ignore = (int) ((float) page.dpi * DEF_BORDER_IGNORE);
// printf("border ignore = %d\n", border_ignore);

PAGE_X1 = border_ignore; // Define page size
PAGE_X2 = page.sizeX - border_ignore;
PAGE_X3 = PAGE_X2 - PAGE_X1;
PAGE_Y1 = border_ignore;
PAGE_Y2 = page.sizeY - border_ignore;
PAGE_Y3 = PAGE_Y2 - PAGE_Y1;

// ***** THIS IS JUST FOR TESTING *****
xs_flush();
printf("Enter 0 to continue? ");
i = auto_input_int();

// clear_scale_and_plot_image(page);

printf("Pausing because page is displayed. Input 0 to segment page? ");
border_ignore = auto_input_int();

while(1) {
    segment_page(page, &clist);
    xs_flush();

    printf("Program finished. Enter 0 to run again? ");
    int ttttemp;
    ttttemp = auto_input_int();
}

charlist_terminate(&clist);
auto_terminate();
}

```


【図241】

```

/* File: util.c
 * This file contains general utility routines.
 *
 * ut_reverse(width, height, buf);
 * ut_getbits(x, p, n);
 * ut_ascii_bitmap(width, height, buf);
 *
 * Usage:
 *
 * Just call the routines.
 */
/* ----- Include ----- */
#include <stdio.h>
#include <stdlib.h>
/* ----- Global ----- */
/* ----- External ----- */
/*
 * Routine: ut_reverse(width, height, buf)
 * Description: Reverse the bits for display format
 * Return: Nothing
 * Author: Rick Lee
 */
/* ut_reverse(width, height, buf)
 * int width;
 * int height;
 * char *buf;
 */
{
    unsigned char val, new_val;
    int x, y;
    for(x = 0; x < (width-1); height, x++)
    {
        x = 0;
        val = buf[x];
        new_val = 0;
        for(y = 1; y >= 0; y--)
        {
            if(ut_getbits(val, x-1))
                new_val |= 0x01 << y;
            new_val |= new_val;
            buf[x] = new_val;
        }
    }
}

/*
 * Routine: ut_getbits(x, p, n)
 * Description: Check if the bit is turned on or not
 * Return: Nothing
 * Author: Rick Lee
 */
ut_getbits(x, p, n)

```

```

unsigned char x; /* The character to allow bits from */
int p; /* The beginning of the bit number */
int n; /* Number of bits to get */
{
    return(x >> (p-n)) & ~(-1 << n);
}

/*
 * Routine: ut_ascii_bitmap(width, height, buf)
 * Description: Display the bitmap using ascii characters
 * Return: Nothing
 * Author: Rick Lee
 */
ut_ascii_bitmap(width, height, buf)
{
    int width;
    int height;
    char *buf;
    {
        short bit_code = 0x0000;
        short bit_val;
        short x, y, i, bits;
        for(x = 0; x < height; x++)
        {
            for(y = 0; y < width; y++)
            {
                bits = buf[y];
                for(i = 0; i < 8; i++)
                {
                    bit_result = bits & bit_code;
                    if(bit_result == 0)
                        printf(" ");
                    else
                        printf("X");
                    bit_code = bit_code >> 1;
                }
                bit_code = 0x0000;
            }
            printf("\n");
        }
    }
}

```


【図243】

```

XCopyArea(XtDisplay(canvas), pix, XtWindow(canvas), line_gc, 0,
width.height, x, y);

/* Now free the bitmap */
XFreePixmap(XtDisplay(canvas), bitmap);

/* Now free the bitmap */
XFreePixmap(XtDisplay(canvas), pix);

/* Now flush the queue to X-server */
XFlush(XtDisplay(canvas));
}

/*
Routine: xs_clr_win()
Description: Clear the screen
Return: Nothing
Author: Rick Lee
*/
xs_clr_win()
{
XClearArea(XtDisplay(canvas), XtWindow(canvas), 0, 0, 0, 0, TRUE);
}

/*
Routine: xs_draw_box(x, y, width, height)
Description: Draw a box on screen
Return: Nothing
Author: Rick Lee
*/
xs_draw_box(x, y, width, height)
int x;
int y;
int width;
int height;
{
XDrawRectangle(XtDisplay(canvas), XtWindow(canvas), line_gc,
x, y, width, height);
}

/*
Routine: xs_string
Description: Draw a string on the screen
Return: Nothing
Author: Rick Lee
*/
xs_string(x, y, str)
int x;
int y;
char *str;
{
XDrawString(XtDisplay(canvas), XtWindow(canvas), line_gc,
x, y, str, strlen(str));
}

/*
Routine: xs_setcolor
Description: Set the current color
Return: Nothing
Author: Rick Lee
*/
xs_setcolor(color)
int color;
{
gcv.foreground = color;
XSetForeground(XtDisplay(canvas), line_gc, color);
}

/*
Routine: xs_win_dim
Description: Get the size of the window
Return: Nothing
Author: Rick Lee
*/
xs_win_dim(width, height)
int *width;
int *height;
{
Arg args[2];
Dimension win_width, win_height;
XSetArg(args[0], Xcwidth, win_width);
XSetArg(args[1], Xcheight, win_height);
XGetValues(canvas, args, 2);

*width = win_width;
*height = win_height;
}

/*
xs_set_map(u)
Widget w;
{
int ncolors;
XColor colors[256];
Colormap def_map;
int count = 1;

Display *d;
int scr;
Colormap def_colormap;
XColor color;
int w;
int the_color = 2080;

d = XtDisplay(w);
scr = DefaultScreen(d);
ncolors = DisplayCells(d, scr);
def_colormap = DefaultColormap(d, scr);
for(x = 0; x < ncolors; x++)
{
color.red.red = w;
}
}

```

—227—

[illegible]

【図247】

```

//
// File: 1270vpp.c
// This routine contains the lines to find blocks that have non-zero vpp
// (vpp: 1 segmentation)
//
// =====
//
#include "charlist.h"
#include "cinaps.h"

// =====
// Scan a line quickly (every 1/4 pixel) looking for an 'on' pixel
int line_scan_quick( // Returns 1 for pixel found, 0 otherwise
cinaps line, // The line to process
int c) // The column to scan
{
    int r;
    for (r=0; r<line.size; r++)
        if (line.pixel(r)(c) & 1)
            return 1;
    return 0;
}

// =====
// Scan a line completely looking for an 'on' pixel
int line_scan_complete( // Returns 1 for pixel found, 0 for otherwise
cinaps line, // The line
int c) // The column to scan
{
    int r;
    for (r=0; r<line.size; r++)
        if (line.pixel(r)(c) & 1)
            return 1;
    return 0;
}

// =====
// The main routine. Call this to make the list
void extract_vpp_list(
cinaps line, // The image of the line to segment
struct NonZeroVppList *list) // A pointer to the list
{
    int c;
    int left, right;
    for (c=0; c<line.size; c++) // Scan the image
    {
        if (line_scan_quick(line, c) // If found something
        {
            left = right = c; // Look to the left for the start
            while (--c >= 0)
                if (line_scan_complete(line, c) == 0)
                    break;
            left = c;
            c = right; // Keep looking to the right
            while (++c < line.size)
                if (line_scan_complete(line, c) == 0)
                    break;
            right = c;
            NonZeroVppAdd(list, left, right); // Add in the new break
        }
    }
}

```

【手続補正書】

【提出日】平成5年5月13日

【手続補正1】

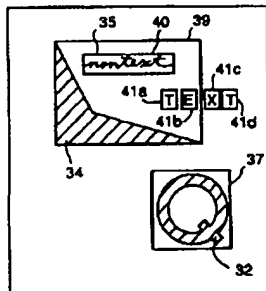
【補正対象書類名】図面

【補正対象項目名】全図

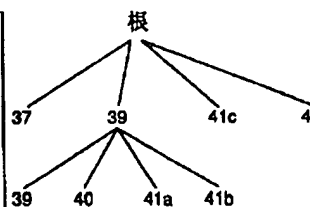
【補正方法】変更

【補正内容】

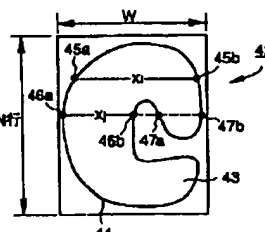
【図5B】



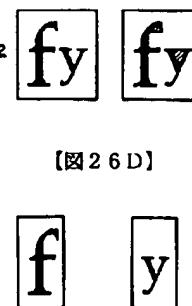
【図5C】



【図6A】



【図26B】【図26C】

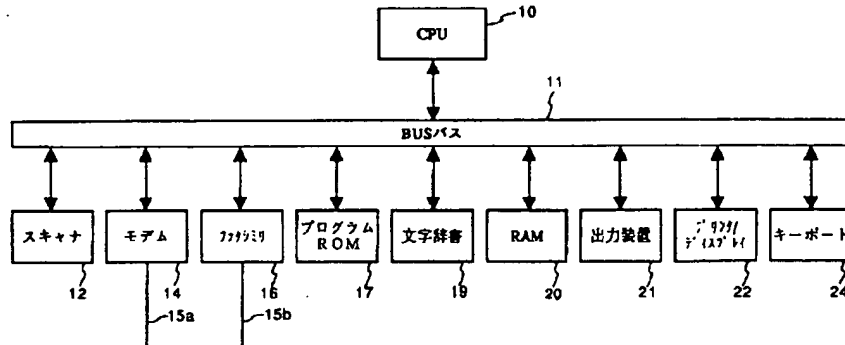


【図26D】

【図1】 * *

【図35】

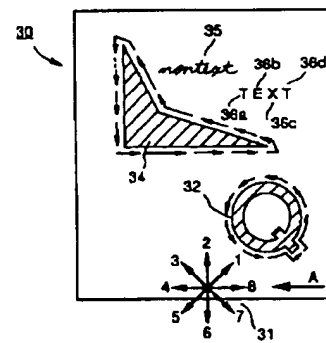
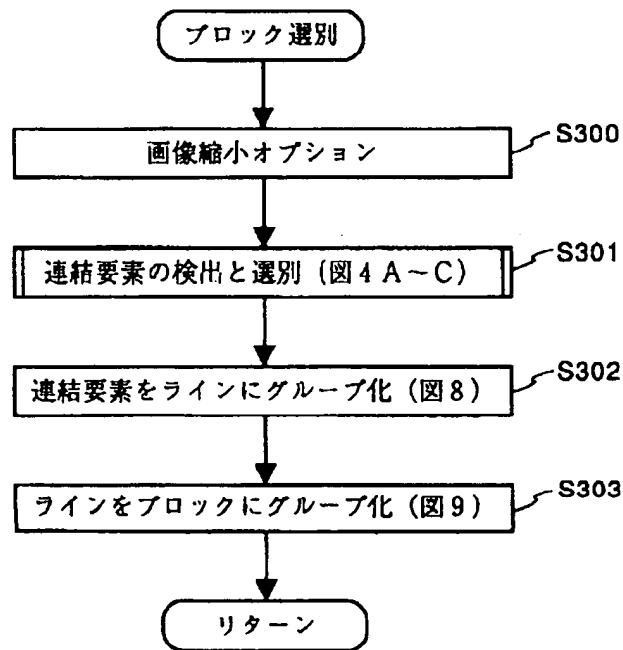
Source Code For
Block Selection



【図3】

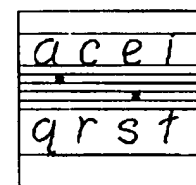
※ ※

【図5A】



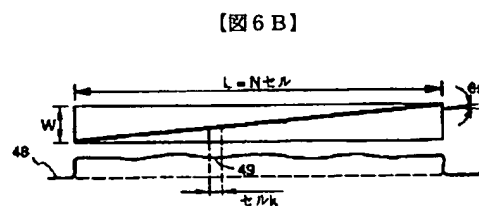
【図18B】

【図234】

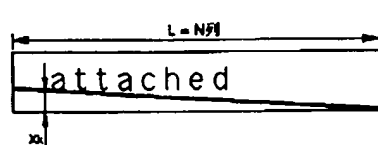


【図6C】

【図194】



【図126】

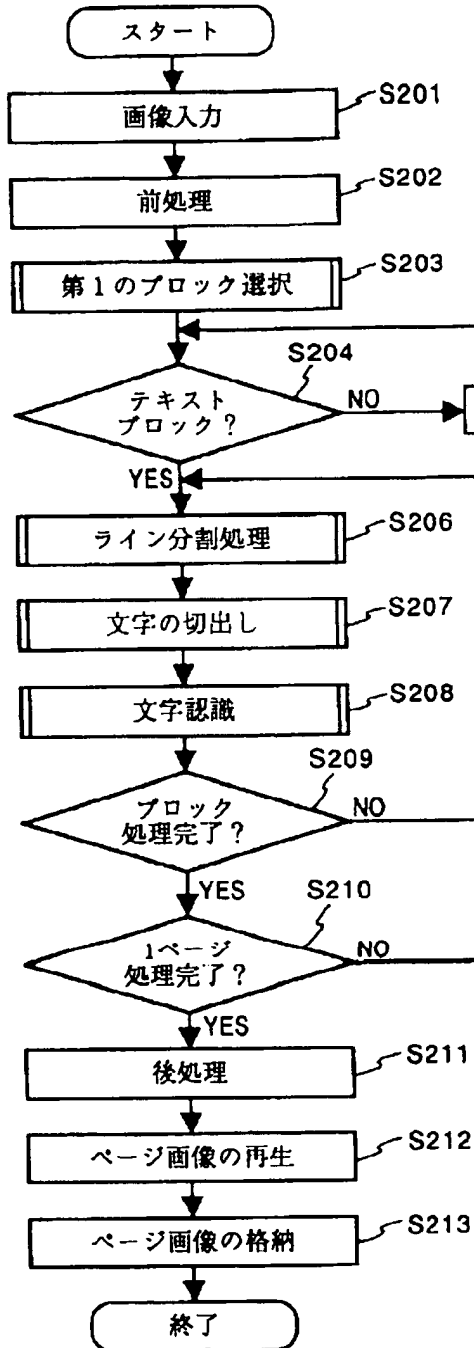


```

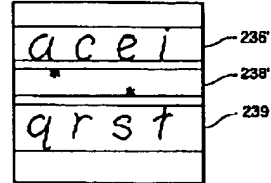
#define FILE_LENGTH      80
#define LINE_LENGTH      256
#define FONT_SIZE_LENGTH 30

```

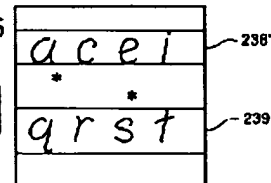
【図2】



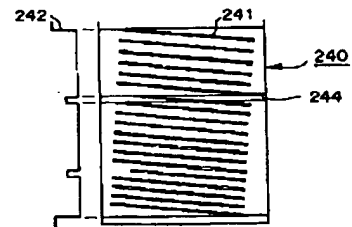
【図18C】



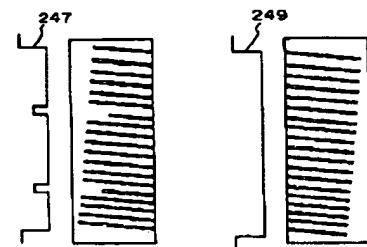
【図18D】



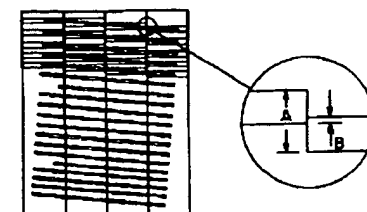
【図19A】



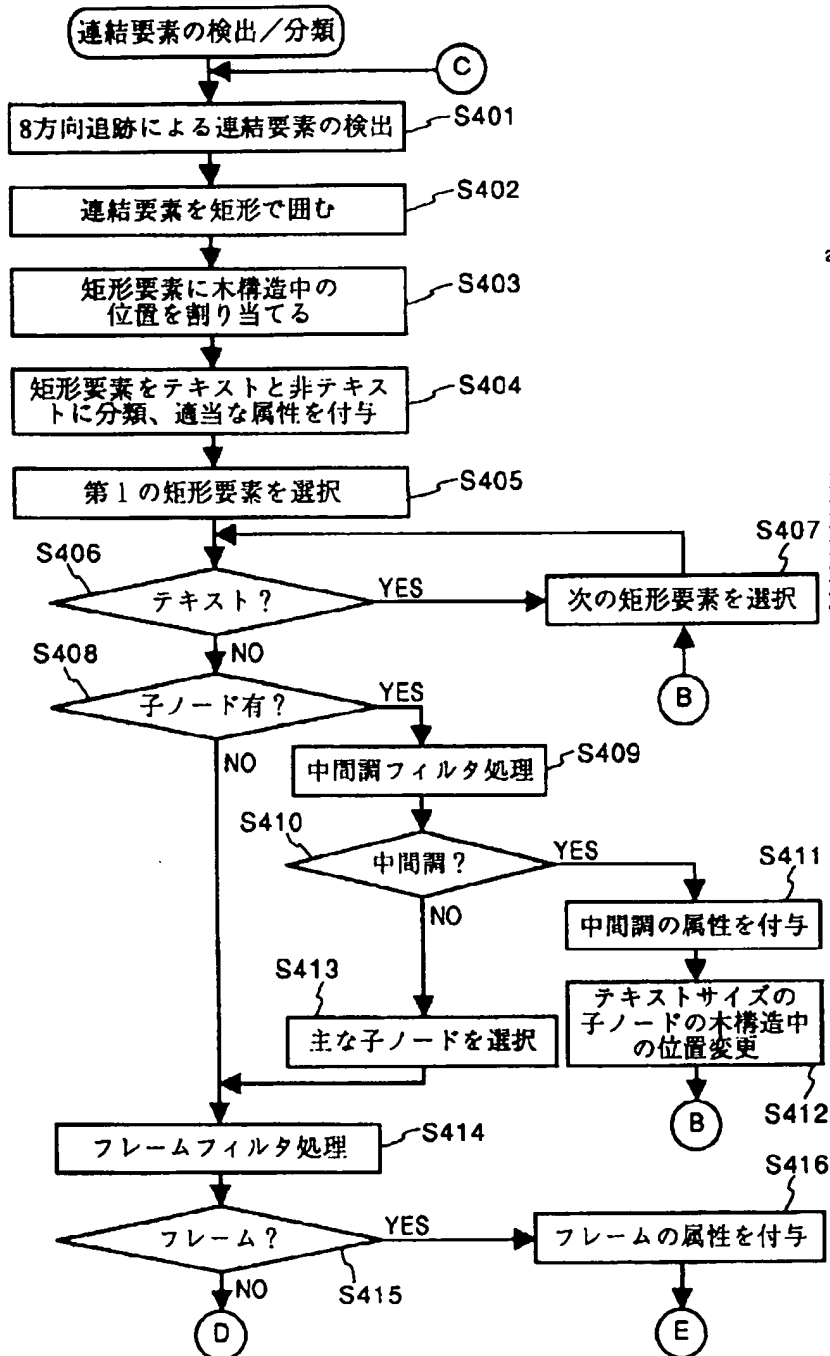
【図19B】



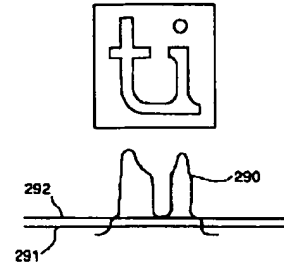
【図19D】



【図4A】



【図29A】



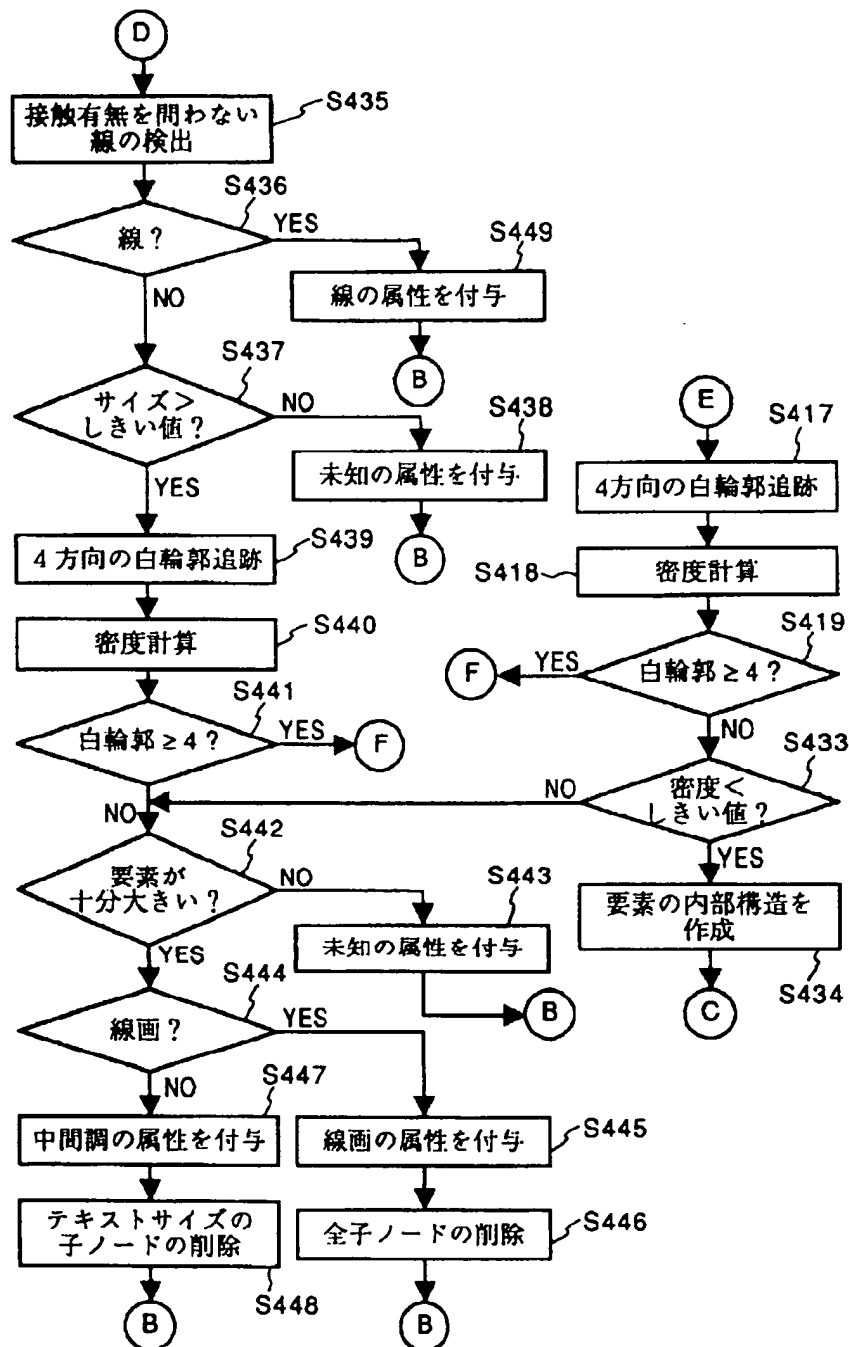
【図134】

```

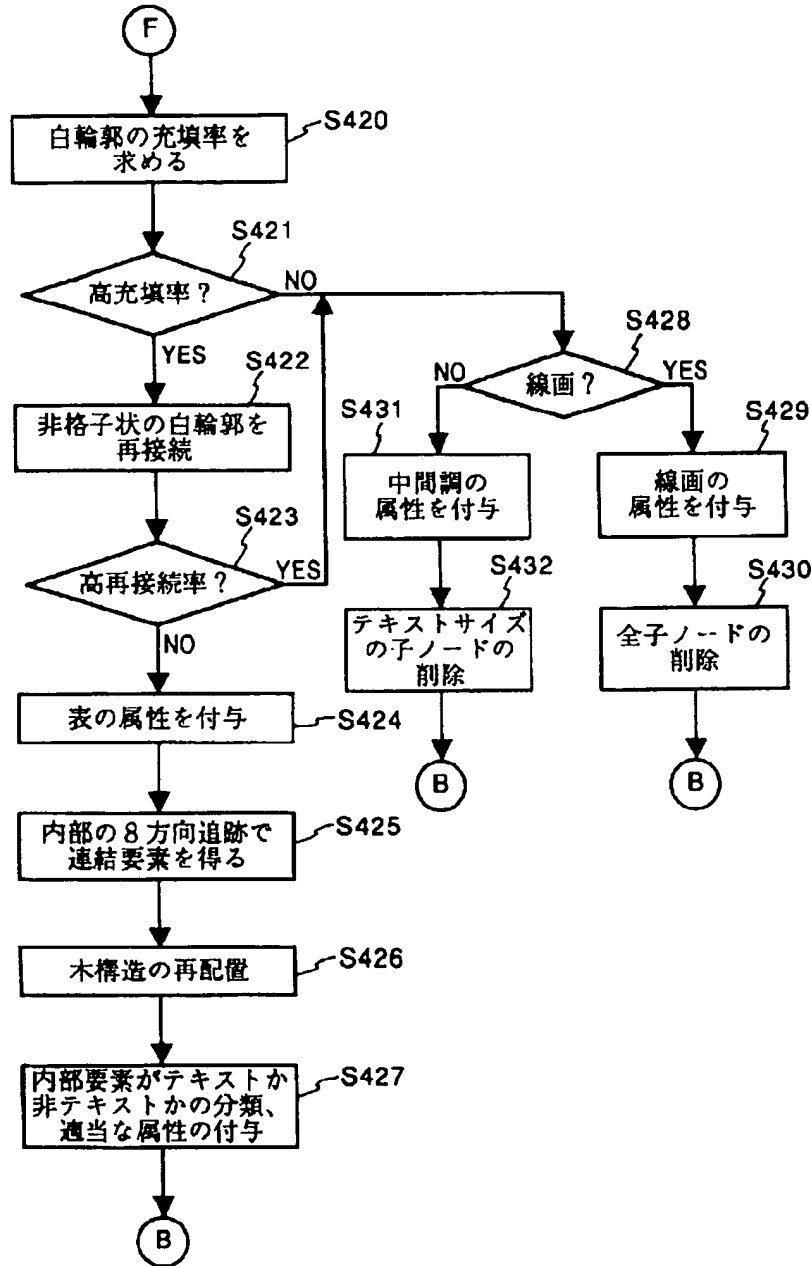
extern float *vector();
extern unsigned char *cvector();
extern int *lvector();
extern double *dvector();
extern unsigned char **cmatrix();
extern float **lmatrix();
extern int **lmatrix();
extern double **dmatrix();
extern void free_vector();
extern void free_lvector();
extern void free_dvector();
extern void free_matrix();
extern void free_lmatrix();
extern void free_dmatrix();

```


【図4B】



【図4C】



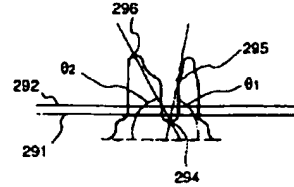
【図26A】

Satisfy

【図34B】

Satisfy

【図29B】



【図34A】



【図163】

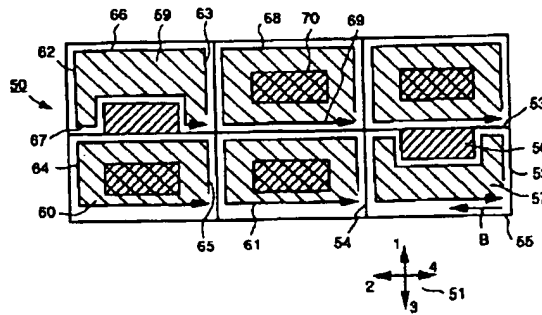
```

HISTOGRAM_THRESHOLD);
charon = charon->next; // Get the next character
}
free_ivector(histogram); // Deallocate the memory

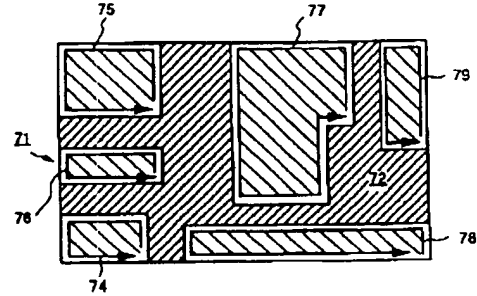
```

【図7A】

* *



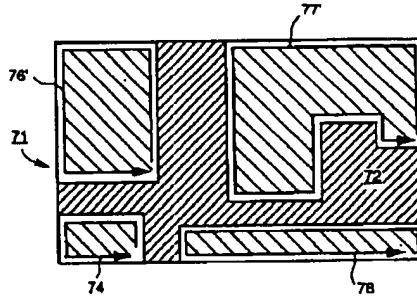
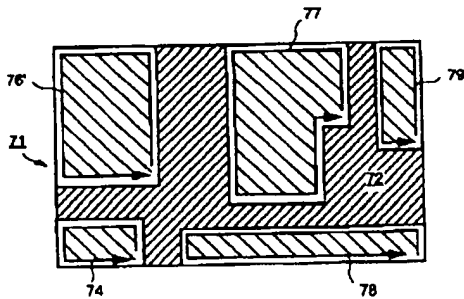
【図7B】



【図7C】

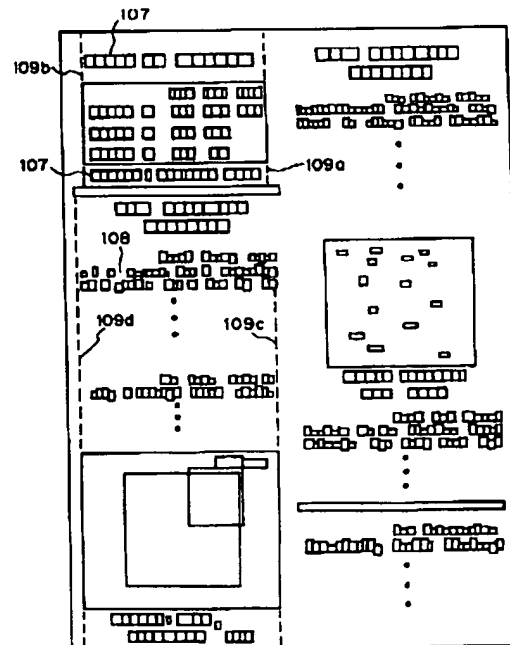
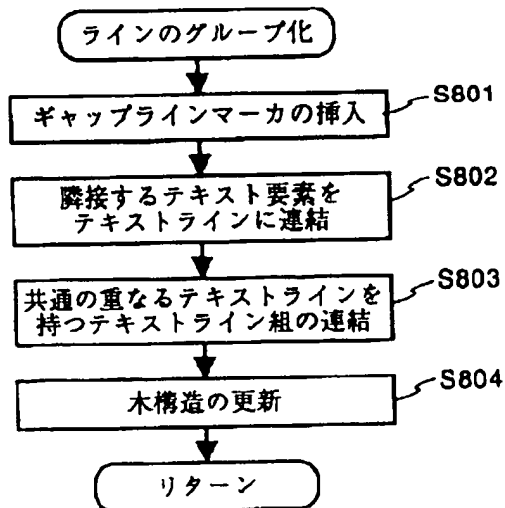
※ ※

【図7D】



【図8】

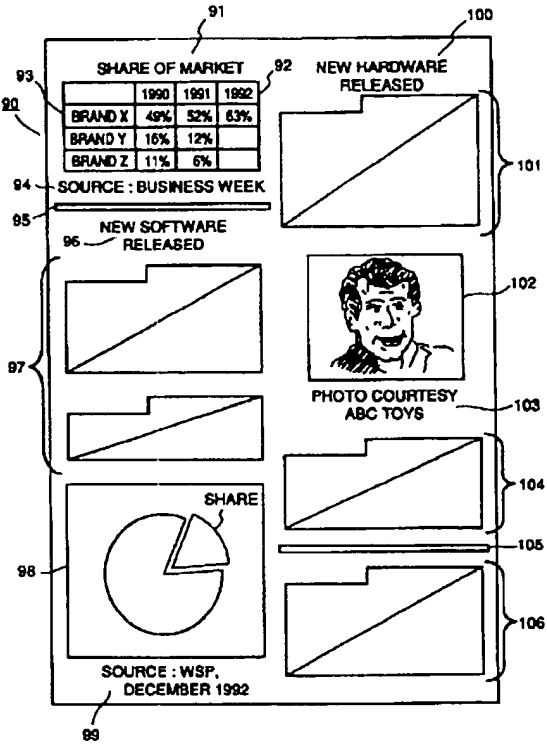
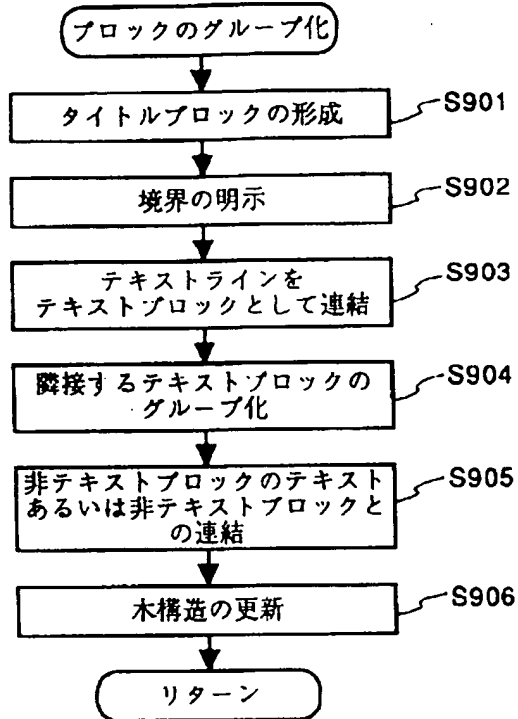
【図11】



【図9】

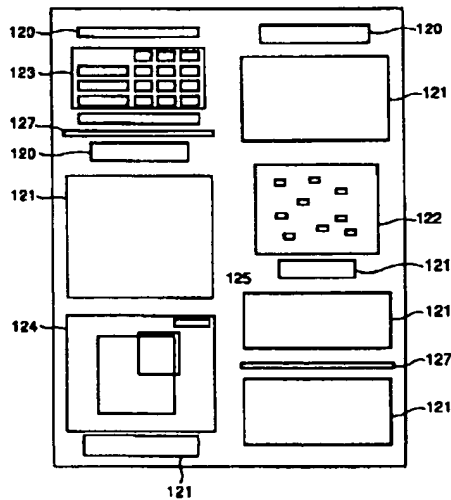
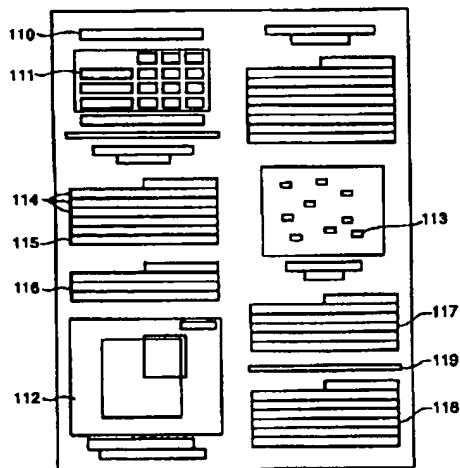
* *

【図10】

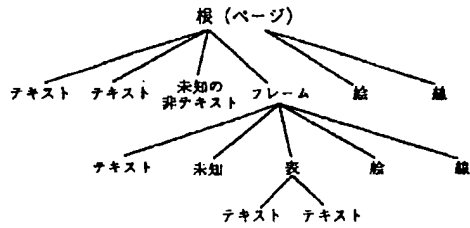


【図12】

【図13】

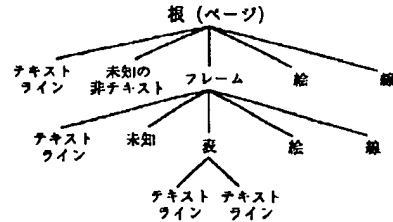


【図14】

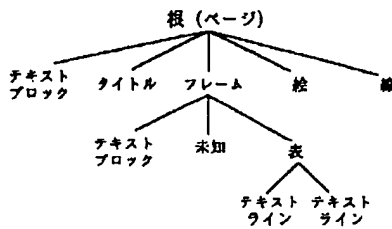


* *

【図15】

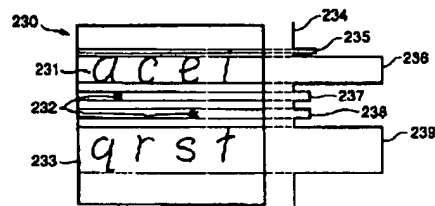


【図16】



※ ※

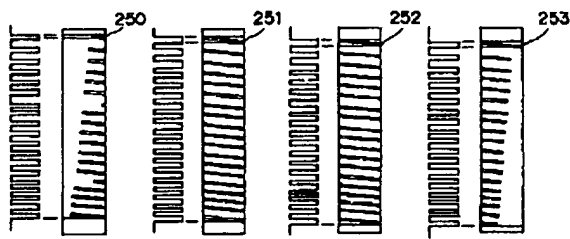
【図18A】



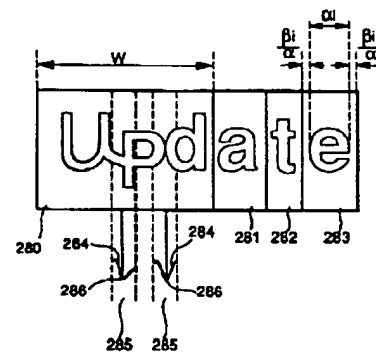
【図19C】

★ ★

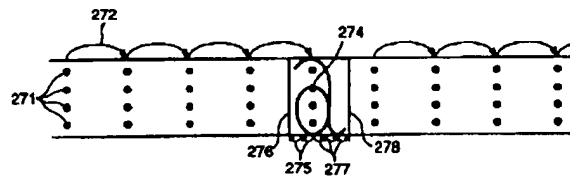
【図24】



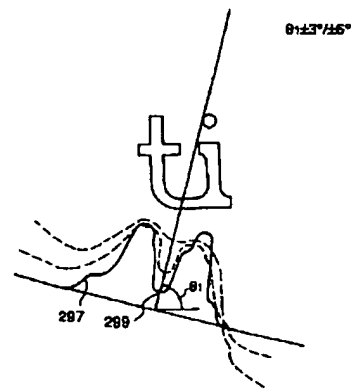
【図22】



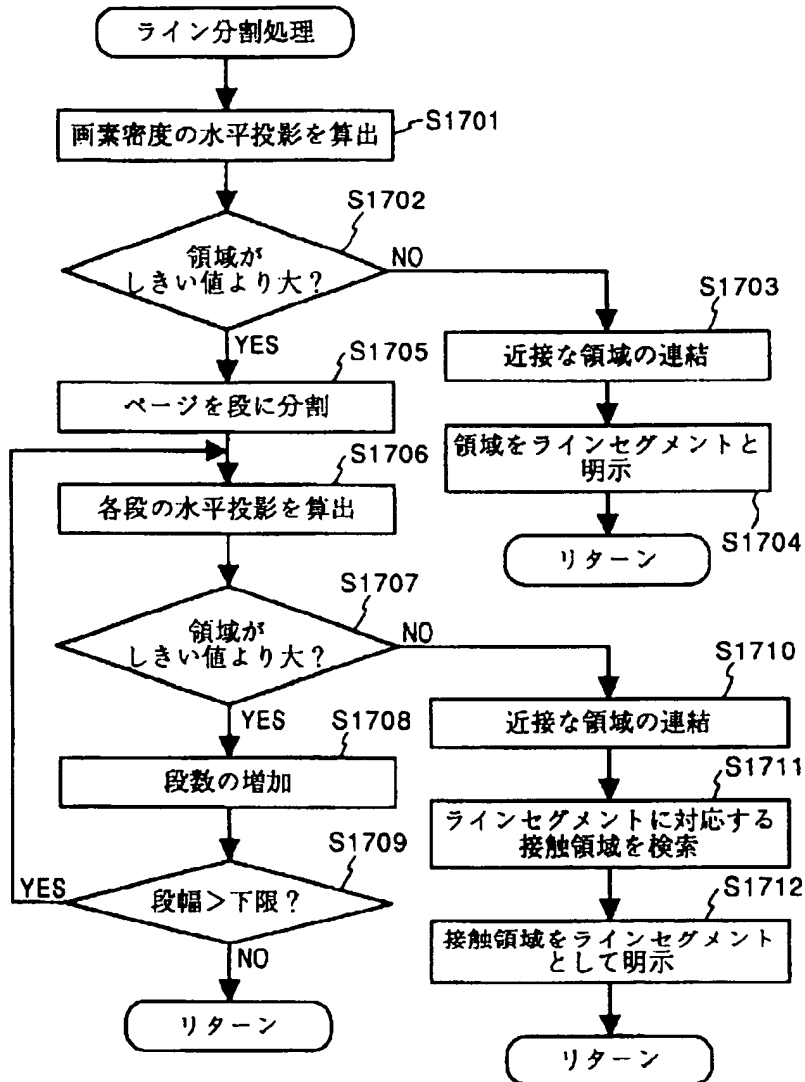
【図29C】



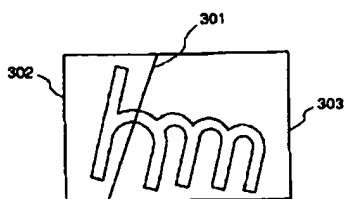
【図29D】



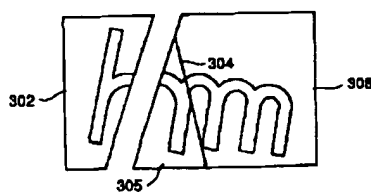
【図17】



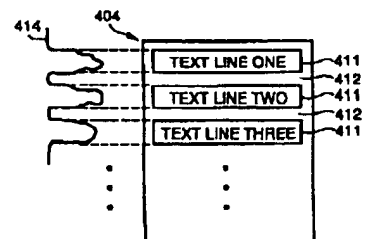
【図31A】



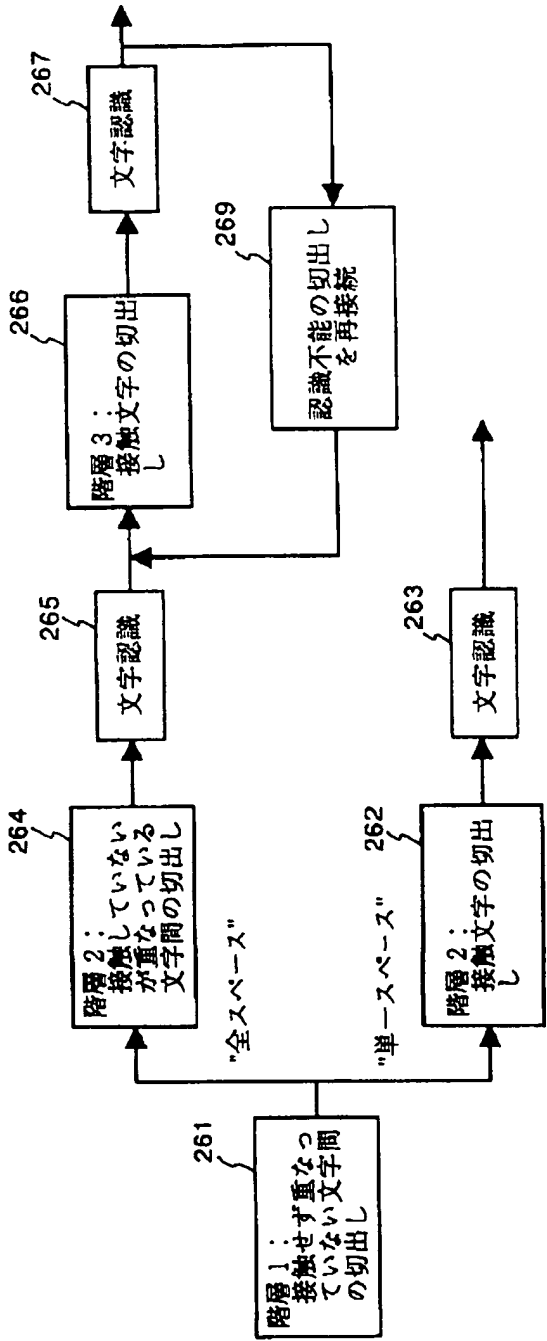
【図31B】



【図33A】



【図20】



【図36】

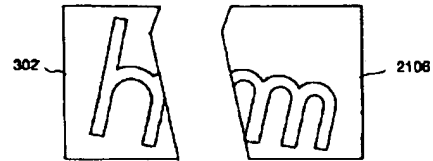
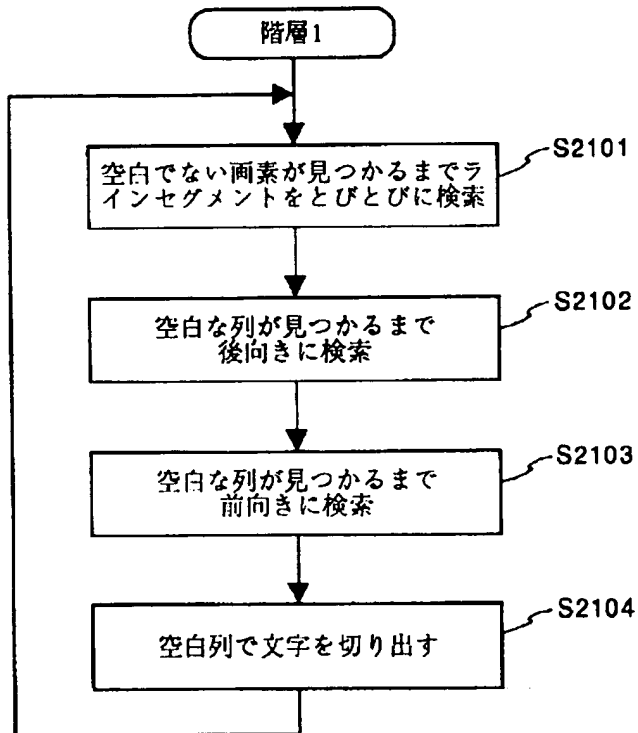


```
#include "util.h"
main(argc,argv)
char *argv;
{
    sv = popup_block();
    /*-----*/
    /* Initialize Architect */
    /*-----*/
    OrtpLevel = OrtpLevel*OCB Block Selection Demo', &argc, argv);
    /* Create and popup the first window of the interface. */
    /*-----*/
    sv = popup_block();
    /* Enter the event loop; this function never returns. */
    /*-----*/
    while(1);
}
```

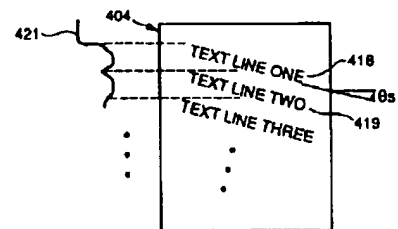
【図21】

* *

【図31C】

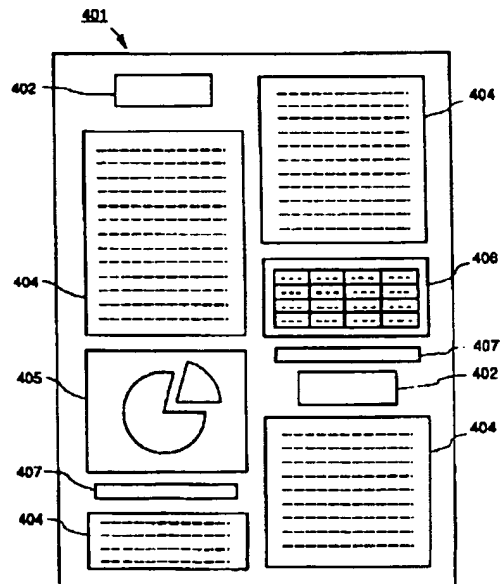
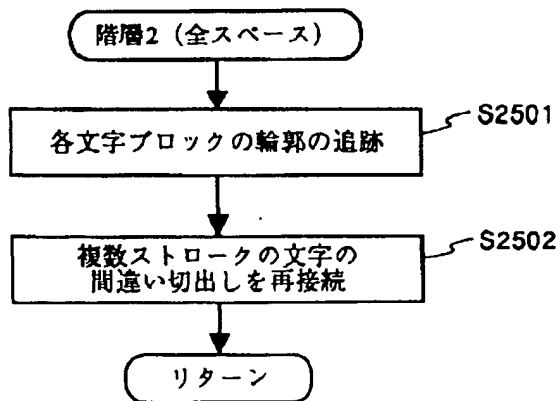


【図33B】



【図25】

【図32】



【図129】

```

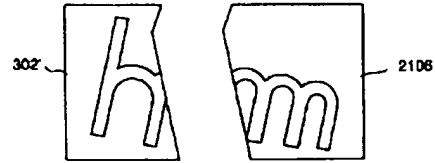
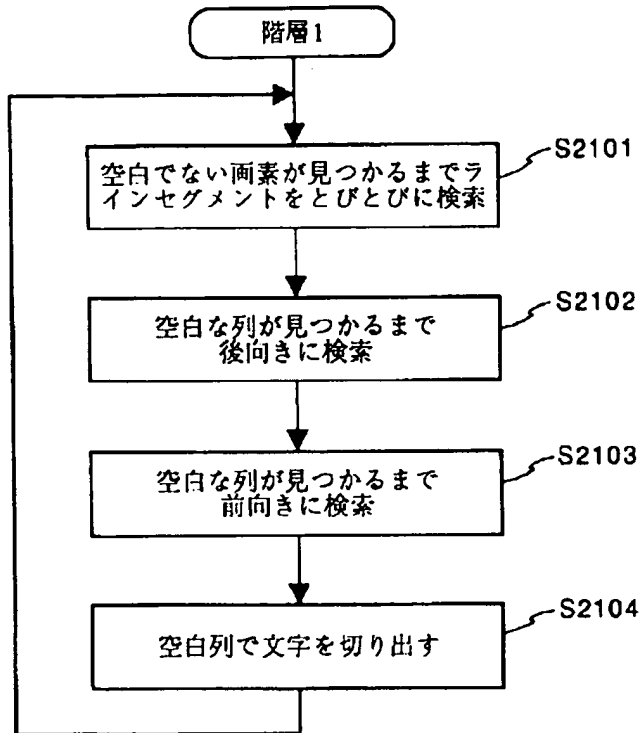
struct sectionblock "previous;
struct sectionblock "next;

int upper_y, lower_y, left_x, right_x;
int width, length;
int index;
int multilineblock;
int att;
int boundary_index;
};
  
```

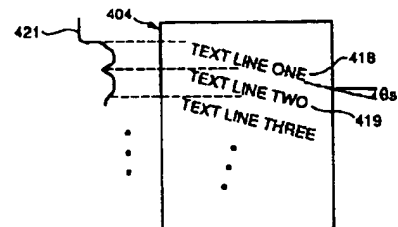

【図21】

* *

【図31C】

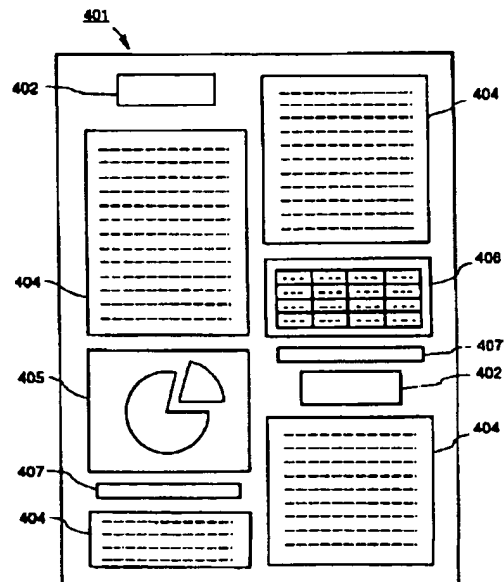
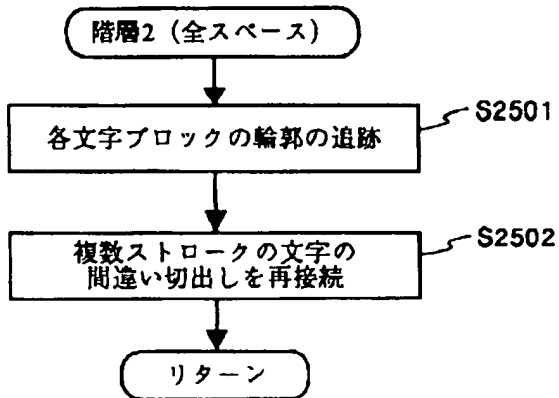


【図33B】



【図25】

【図32】

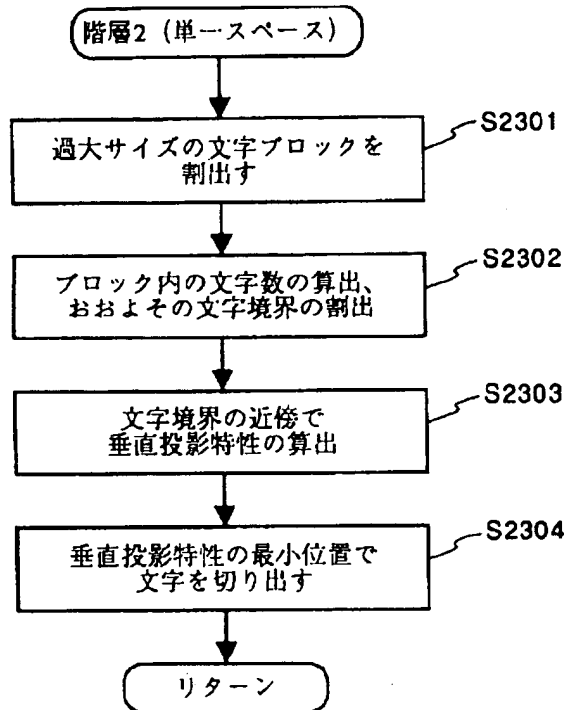


【図129】

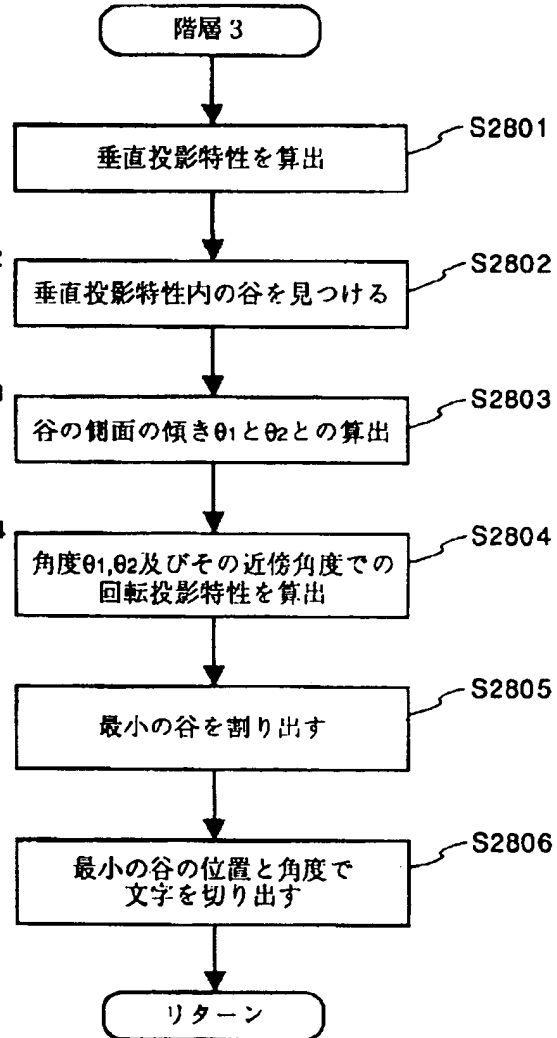
```

struct sectionblock "previous;
struct sectionblock "next;
int upper_y, lower_y, left_x, right_x;
int width, length;
int index;
int runlineblock;
int att;
int boundary_index;
};
  
```

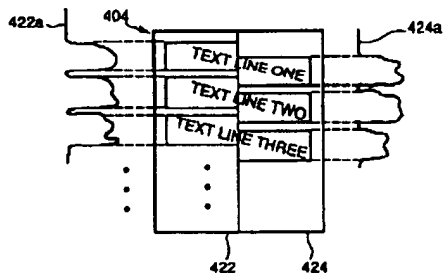
【図23】



【図28】



【図33C】



【図14】

```

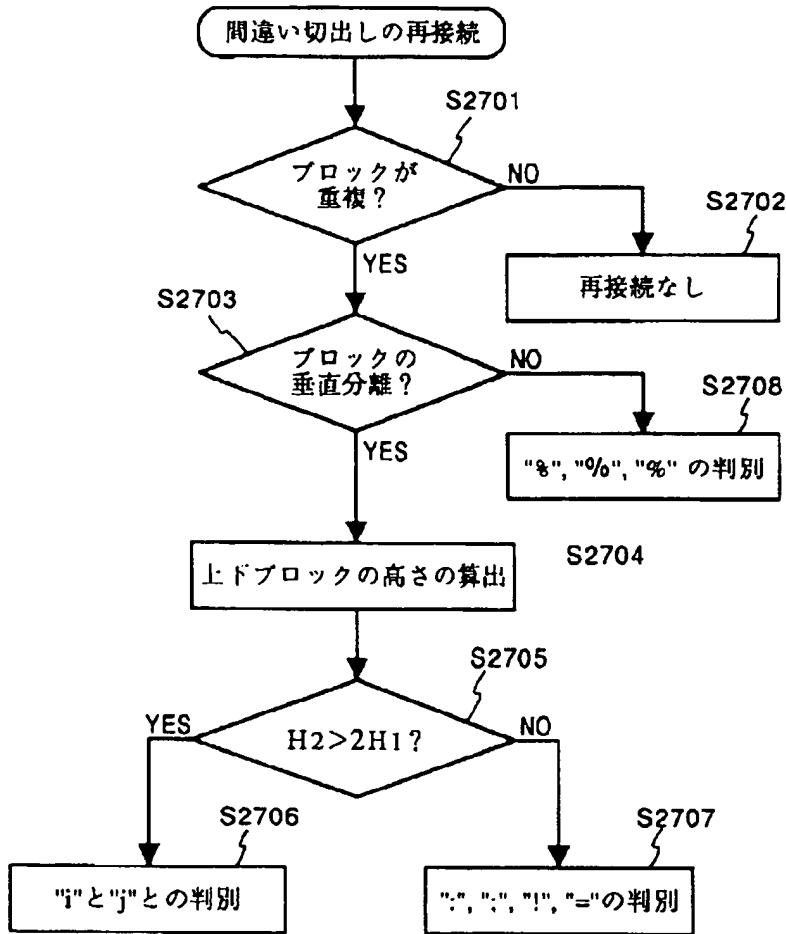
/* ARGSUSED */
static void okCallback_QuitMessageBox(UiWidget, UiContext, UiCallbackArg)
{
    UiWidget      UiWidget;
    _UiQuitDialog _UiQuitDialog;
    int           _UiCallbackArg;

    UiWidget      UiThisWidget;
    UiThisWidget= UiWidgetToUiWidget (UiWidget);
    _UiFromContext (UiContext);

    {
        exit (1);
    }

    _UiToContext (UiContext);
}
  
```

【図27】



【図46】

```

/*
 * _UnCon->DisplayForm= DisplayForm;
 * _UnCon->DisplayFormClose_Func= DisplayFormClose_Func;
 * _UnCon->DisplayForm_DrawArea= DisplayForm_DrawArea;
 */
static void _UnFromContext(_UnCon)
{
    _UnCon->DisplayForm = _UnCon;
    DisplayForm= _UnCon->DisplayForm;
    DisplayFormClose_Func= _UnCon->DisplayFormClose_Func;
    DisplayForm_DrawArea= _UnCon->DisplayForm_DrawArea;
}

/* ARGUSED */
static void activateCallback_DisplayFormClose_Func(UiWidget, UiContext, UiCallbackArg)
{
    UiWidget _UnWidget;
    _UnConDisplayForm _UnContext;
    int _UnCallbackArg;
    {
        sWidget _UnThisWidget;
        _UnThisWidget= _UnWidgetToUiWidget(_UnWidget);
        _UnFromContext(_UnContext);
        {
            _UnPopdownInterface(DisplayForm);
        }
        _UnToContext(_UnContext);
    }
}

```

【図133】

```

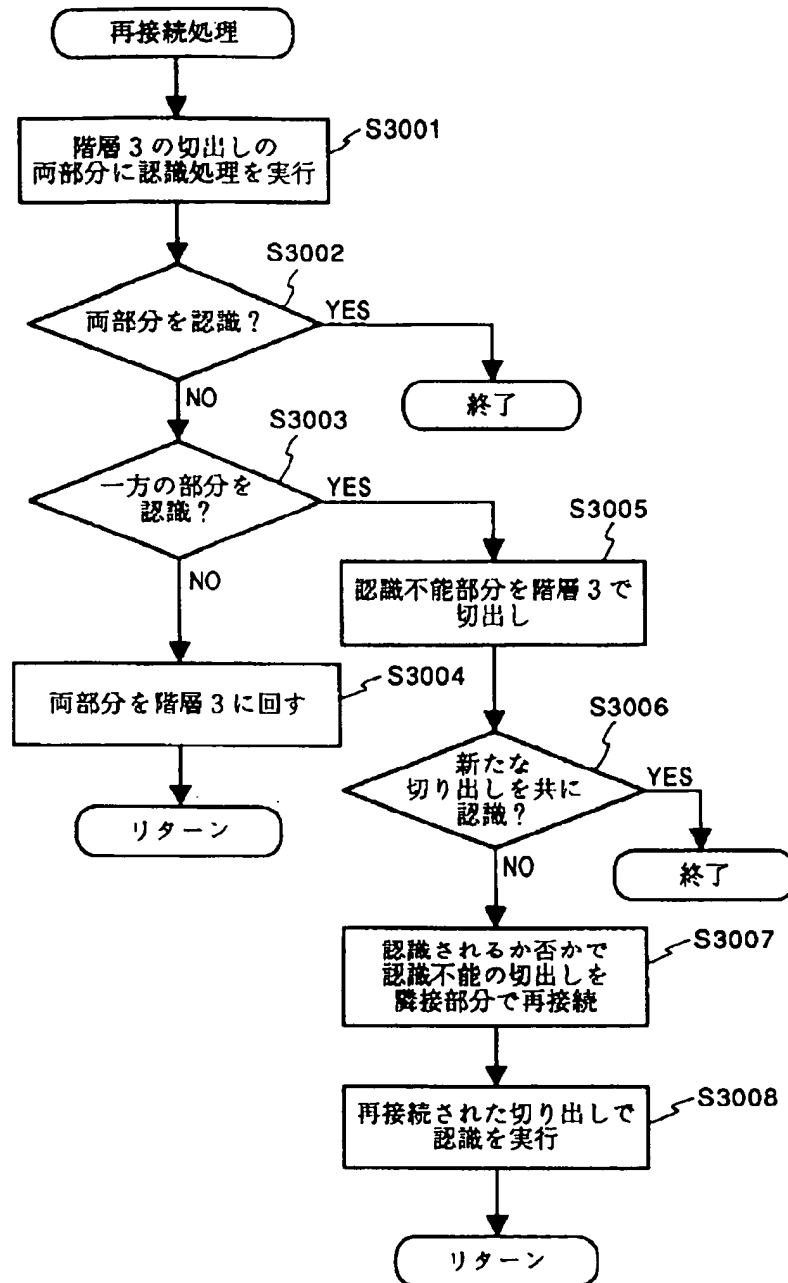
#define X 0
#define G 1
#define B 2

int color_comp[9][3]={{0x65, 0xcd, 0x00},
                      {0xf0, 0x80, 0x00},
                      {0xf0, 0x00, 0xf0},
                      {0xf0, 0xf0, 0x00},
                      {0xf0, 0xf0, 0x00},
                      {0x98, 0x00, 0x00},
                      {0x65, 0xf0, 0xf0},
                      {0x65, 0x00, 0xf0},
                      {0x00, 0xf0, 0x00}};

/*
int color_comp[9][3]={{0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00},
                      {0x00, 0x00, 0x00}};
*/

```

【図30】



【例 37】

[illegible]

—245—

[illegible]

—246—

[illegible]

[illegible]

【図42】

```

_UxCon->FileSelectionDialog= FileSelectionDialog;
_UxCon->FileSelectionBox= FileSelectionBox;
_UxCon->parent= parent;
}
static void _UxFromContext(_UxCon)
_UxCFileSelectionDialog *_UxCon;
{
    FileSelectionDialog= _UxCon->FileSelectionDialog;
    FileSelectionBox= _UxCon->FileSelectionBox;
    parent= _UxCon->parent;
}
/* ARGSUSED */
static void cancelCallback_FileSelectionBox(UxWidget, UxContext, UxCallbackArg)
Widget UxWidget;
_UxCFileSelectionDialog *_UxContext;
int UxCallbackArg;
{
    sWidget UxThisWidget;
    UxThisWidget= UxWidgetToWidget(UxWidget);
    _UxFromContext(UxContext);
    {
        UxPopdownInterface(FileSelectionDialog);
    }
    _UxToContext(UxContext);
}
/* ARGSUSED */
static void okCallback_FileSelectionBox(UxWidget, UxContext, UxCallbackArg)
Widget UxWidget;
_UxCFileSelectionDialog *_UxContext;
int UxCallbackArg;
{
    sWidget UxThisWidget;
    UxThisWidget= UxWidgetToWidget(UxWidget);
    _UxFromContext(UxContext);
    {
        handle_tiff_handle;
        UxPopdownInterface(FileSelectionDialog);
        strcpy(filename, get_input_filename());
        printf("Open (%s)\n", filename);
        tiff_handle = UxCreateSubproc("/usr/local/bin/X11/imageview", filename,
UxAppendTo);
        if (tiff_handle == -1) {
            error_create_subproc();
            return;
        }
        if (UxRunSubproc(tiff_handle, NULL) == -1) {
            error_run_subproc();
            return;
        }
        open_tiff_file(filename);
    }
    _UxToContext(UxContext);
}
}

```

【図48】

```

}
static void _UxFromContext(_UxCon)
_UxCInfoForm *_UxCon;
{
    InfoForm= _UxCon->InfoForm;
    InfoFormClose_Pkg= _UxCon->InfoFormClose_Pkg;
    InfoFormArea= _UxCon->InfoFormArea;
}
/* ARGSUSED */
static void activateCallback_InfoFormClose_Pkg(UxWidget, UxContext, UxCallbackArg)
Widget UxWidget;
_UxCInfoForm *_UxContext;
int UxCallbackArg;
{
    sWidget UxThisWidget;
    UxThisWidget= UxWidgetToWidget(UxWidget);
    _UxFromContext(UxContext);
    {
        UxPopdownInterface(InfoForm);
    }
    _UxToContext(UxContext);
}
}

```

【図62】

```

int width, height;
{
    int x, y, xdis, ydis, i;
    static fid = 33;
    init_graphics(wd);
    xs_clearwin(wd);
    xdis = ((height-30)/(TOTALCOLOR));
    xs_getfont(wd);
    for (i = 0; i < TOTALCOLOR; i++){
        xs_set_foreground(color_code[i], wd);
        xs_setfont(wd, fid);
        xs_string(wd, 30, (i+1)*xdis+30, color_label[i]);
    }
    xs_flush(wd);
}

```

[illegible]

【例 49】

```

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <unistd.h>
#include <sys/types.h>
#include <sys/stat.h>
#include <fcntl.h>
#include <time.h>
#include <sys/time.h>

struct timeval first, second, lapsec;
struct timespec ttp;

/* Start the timing clock */
start_clock ()
{
    gettimeofday (&first, &ttp);
}

/* Stop the timing clock and return the elapsed time in milliseconds. */
double stop_clock ()
{
    double time;
    gettimeofday (&second, &ttp);
    if (first.tv_usec > second.tv_usec) {
        second.tv_usec -= 1000000;
        second.tv_usec += first.tv_usec;
    }
    lapsec.tv_usec = second.tv_usec - first.tv_usec;
    lapsec.tv_sec = second.tv_sec - first.tv_sec;
    time = (double)((double)lapsec.tv_usec / (double)1000) +
        (double)lapsec.tv_sec;
    return time;
}

/* Return the basename of a file without the directory. */
char *extract_base_file_no_dir (name)
    char *name;
{
    char tap(FILE_LENGTH), *pos;
    strcpy (tap, name);
    if ((pos = strchr(tap, '/')) != NULL)
        return pos+1;
    else
        return name;
}

char *extract_base_file (name)
    char *name;
{
    char tap(FILE_LENGTH), *pos;
    strcpy (tap, name);
    if ((pos = strchr(tap, '/')) != NULL)
        return pos+1;
    else
        return name;
}

/* Return the basename of a file */
char *extract_base_file (name)
    char *name;
{
    char tap(FILE_LENGTH), *pos;
    strcpy (tap, name);
    if ((pos = strchr(tap, '/')) != NULL)
        return pos+1;
    else
        return name;
}

/* Return -1 if file cannot open */
int check_file (name)
    char *name;
{
    FILE *fp;
    if ((fp = fopen (name, "r")) == NULL)
        return -1;
    fclose (fp);
    return 1;
}

/* Error message */
char *msg;
print_error (msg)
{
    printf (":Error: %s\n", msg);
}

/* Create subprocess */
error_create_subproc ()
{
    printf (":Error: Cannot create a subprocess.\n");
}

/* Exit callback */
error_subproc_exit_callback ()
{
    printf (":Error: Cannot set subprocess exit callback.\n");
}

/* Subprocess */
error_subproc ()
{
    printf (":Error: Cannot start a subprocess.\n");
}

```

—253—

【図51】

```

        } height = win_height;
    }

    xs_draw_line(wd, x1, y1, x2, y2);
    widget wd;
    int x1, y1, x2, y2;
    } xs_draw_line(xtDisplay(wd), xsWindow(wd), drawing_gc, x1, y1, x2, y2);
}

xs_color_paint(wd, x, y, color)
    widget wd;
    int x, y, color;
{
    Arg args[1];
    int foreground;
    XSetArg(args[0], XIntForeground, foreground);
    XSetValues(wd, args, 1);
    xsSetForeground(xtDisplay(wd), drawing_gc, color);
    xsDrawPoint(xtDisplay(wd), xsWindow(wd), drawing_gc, x, y);
    xsSetForeground(xtDisplay(wd), drawing_gc, foreground);
}

xs_color_line(wd, x1, y1, x2, y2, color)
    widget wd;
    int x1, y1, x2, y2, color;
{
    Arg args[1];
    int foreground;
    XSetArg(args[0], XIntForeground, foreground);
    XSetValues(wd, args, 1);
    xsSetForeground(xtDisplay(wd), drawing_gc, color);
    xsDrawLine(xtDisplay(wd), xsWindow(wd), drawing_gc, x1, y1, x2, y2);
    xsSetForeground(xtDisplay(wd), drawing_gc, foreground);
}

xs_color_box(wd, x1, y1, x2, y2, color)
    widget wd;
    int x1, y1, x2, y2, color;
{
    Arg args[1];
    int foreground;
    XSetArg(args[0], XIntForeground, foreground);
    XSetValues(wd, args, 1);
    xsSetForeground(xtDisplay(wd), drawing_gc, color);
    xsDrawRectangle(xtDisplay(wd), xsWindow(wd), drawing_gc, x1, y1, x2, y2);
    xsSetForeground(xtDisplay(wd), drawing_gc, foreground);
}

xs_color_area(wd, x, y, str, color)
    widget wd;
    int x, y, color;
    char *str;
{
    Arg args[1];
    int foreground;
    XSetArg(args[0], XIntForeground, foreground);
    XSetValues(wd, args, 1);

```

```

        xsSetForeground(xtDisplay(wd), drawing_gc, color);
        xsSetString(xtDisplay(wd), xsWindow(wd), drawing_gc, x, y, str, strlen(str));
        xsSetForeground(xtDisplay(wd), drawing_gc, foreground);
    }

    static char **font;
    XFontSet font(wd);
    widget wd;
    {
        int i, num = 59;
        font = XListFonts(xtDisplay(wd), "...", 59, &num);
        printf("num = %d\n", num);
    }
    xsSetFont(wd, font);
    widget wd;
    {
        XFontStruct *font_info;
        if ((font_info = XQueryFont(xtDisplay(wd), font(fid))) == NULL) {
            printf("num not open the font fid = %d\n", fid);
            exit(-1);
        }
        xsSetFont(xtDisplay(wd), drawing_gc, font_info->fid);
        printf("fid = %d\n", fid);
    }
}

```



```

/*
int getbyte(c, file_position, eoferror)
FILE *f;
long int *file_position;
int *eoferror;

{
    int c;
    c = fgetc(f);
    if (f == EOF) { *eoferror = 1;
        printf("\n Error, Premature End of File :in getbyte()\n\n");
        exit(1);
    }
    (*file_position)++;
    return c;
}

/* Read in an integer from a file. (USD, MSB)
long int getbytes(c, fp)
FILE *f;
long int *fp;

{
    long int c1;
    long int c2;
    c1 = getbyte(f, fp, 1);
    c2 = getbyte(f, fp, 1);
    c3 = (long int)(c1 * 256);
    return c3;
}

/* Read in a long integer
long int getbytes(l, fp)
FILE *f;
long int *lp;

{
    long int c1;
    long int c2;
    long int c3;
    long int c4;
    c1 = getbyte(f, fp, 1);
    c2 = getbyte(f, fp, 1);
    c3 = getbyte(f, fp, 1);
    c4 = getbyte(f, fp, 1);
    return c1 * 256 + c2 * 256 + c3 * 256 + c4;
}

/* Read in part of a Tiff file
.....
struct images *readtiff(filename)
char *filename;

{
    FILE *tagfile;
    long int tagfile_position;
    long int width;
    long int height;
    long int black;
    long int bps;
    long int raster_offset;
    long int nuntage, tag, type, len, value; /* tag values
    long int length, all_width;
    int v;
    struct images *img;

    tagfile = fopen(filename, "rb");
    if (!tagfile) { printf("Error: %s\n", filename); return 0; }
}

```

【図53】

```

        printf("\nError in opening %s\n", filename);
        exit(1);
    }
    tagfile_position = 0;
    /* Read in header and make sure it's Intel! */
    i = getbytes(tagfile, &tagfile_position);
    printf("tagfile is %s\n", i);
    if (i != "tag000")
    {
        printf("... Error: TIFF file not in Intel Format!\n");
        exit(1);
    }
    /* Read in Magic number
    i = getbytes(tagfile, &tagfile_position);
    printf("i = %s\n", i);
    if (i != "tag001A")
    {
        printf("... BAD Magic number for TIFF file!\n");
        exit(1);
    }
    /* Find First File directory
    i = getbytes(tagfile, &tagfile_position);
    if (i != "tag002")
    {
        i = getbytes(tagfile, &tagfile_position);
    }
    /* Then in Directory
    numtags = getbytes(tagfile, &tagfile_position);
    printf("numtags = %d\n", numtags);
    while (--numtags > 0)
    {
        tag = getbytes(tagfile, &tagfile_position);
        type = getbytes(tagfile, &tagfile_position);
        value = getbytes(tagfile, &tagfile_position);
        if (type != 1) && ((tag & 32768) == 0)
        {
            printf("Error: Tag Length is 1!\n");
            exit(1);
        }
        printf("tag = %s\n", tag);
        switch (tag)
        {
            case 0x0005: /* NewSubFileType */
            {
                printf("Can't Deal with Non-Standard NewSubFileType\n");
                exit(1);
            }
            break;
            case 0x0007: /* NewSubFileType */
            {
                if (value != 1)
                {
                    printf("Subfile Type not full resolution\n");
                    exit(1);
                }
            }
            break;
            case 0x0100: /* Width */
            {
                width = value;
                break;
            }
            case 0x0101: /* Length */
            {
                length = value;
                break;
            }
        }
    }
    case 0x0102: /* Bits per sample */
    {
        if (value != 1)
        {
            printf("Error: Bits/sample must be 1!\n");
            exit(1);
        }
        break;
    }
    case 0x0103: /* Compression */
    {
        if (value != 1)
        {
            printf("Compression Used! Error!\n");
            exit(1);
        }
        break;
    }
    case 0x0104: /* Black is */
    {
        black = value;
        break;
    }
    case 0x0107: /* Thresholding tag */
    {
        if (value > 3)
        {
            printf("Thresholding Tag = Illegal value!\n");
            exit(1);
        }
        break;
    }
    case 0x0117: /* Offset to Master Data */
    {
        master_offset = value;
        break;
    }
    case 0x0118: /* Bits = value;
    break;
    case 0x0119: /* Bits = value;
    break;
    case 0x011b: /* Bits = value;
    break;
    case 0x0118: /* Bits = value;
    break;
    default:
    {
        if (value & 32768) == 0)
        {
            printf("Warning: An unsupported tag appears in the TIFF file!\n");
        }
        break;
    }
    }
    }
    val = tagfile_position + master_offset;
    i = getbytes(tagfile, &tagfile_position);
    /*----- READ IN FILE -----*/
    all_length = (int) ((length+scale_ratio)/scale_ratio)*scale_ratio;
    all_width = (int) ((width+scale_ratio)/scale_ratio)*scale_ratio;
    printf("all_length = %d\n", all_length);
    printf("all_width = %d\n", all_width);
    /*
    page = image_allocate((int) all_length, (int) all_width);
    printf("page-size = %d\n", page-size);
    printf("black = %d\n", black);
    for (i = 0; i < all_length; i++)
    {
        for (j = 0; j < all_width; j++)
        {
            if ((i+j) == 0)
            {
                v = getbytes(tagfile, &tagfile_position);
            }
        }
    }

```

【図54】

```

if (v == EOF)
{
    printf("\n Error: Premature End of File found in tiff file\n");
    exit(1);
}
if ((v & 128) == black*128)
    page->pixel[i][j] = 0;
else
    page->pixel[i][j] = 1;
v = v << 1;
}
fclose(tagfile);
return PAGE;
}

```

【図86】

```

        release_block(tmp1);
        tmp1 = tmp11;
    }
    for (; tmp1 != NULL; ) {
        tmp12 = tmp1->next;
        release_block(tmp1);
        tmp1 = tmp12;
    }
    free_rectblock(initial_block);
    return;
}
/* Draw text components for display */
void block_circumulate(p, width, height, p_ori)
unsigned char **p;
int width, height;
unsigned char **p_ori;
{
    int i, j;
    /* open file for storing block data */
    original_p = p_ori;
    reduce_factor = SCALE_RATIO;
    for (i = 0; i <= H; i++)
        for (j = 0; j <= W; j++)
            direct_status[i][j] = 0;
    block_file = fopen("block_file.dat", "w");

    text_length = 130.0/SCALE_RATIO;
    total_text = 0;

    /* the root block stands for the whole page image unit */
    root_rectblock = new_block(0,0);
    root_rectblock->width = width;
    root_rectblock->length = height;
    root_rectblock->upper_y = 0;
    root_rectblock->lower_y = height-1;
    root_rectblock->left_x = 0;
    root_rectblock->right_x = width-1;

    /* find blocks inside the page range */
    block_in_range(p, 0, 0, width-1, height-1, root_rectblock);
    printf("get out of block_in_range\n");
    rearrange_rectblock(root_rectblock);
    printf("get out of rearrange_rectblock\n");

    text_discriminate(root_rectblock);

    printf("get out of text_dis\n");

    firstlevelnontextsearch(p, root_rectblock);

    printf("get out of firstlevelnontextsearch\n");
    print_block(1, root_rectblock);

    fclose(block_file);
    printf("before get out of block_construct\n");
}

```

【図137】

Source Code for Proportional Spaced Segmentation

—258—

[illegible]

—259—

[illegible]

【图 5-7】

[illegible]

【図58】

```

for (tapt1 = firstblock->firstwhite, tapt1 != NULL; ) {
    /* tapt1->numwhite; */
    tapt2 = tapt1->numwhite;
    if (tapt2->numwhite != 0) {
        for (i = 0; i < tapt2->numwhite; i++) {
            free_ptr(tapt1->subwhite[i].ptr);
            tapt1->subwhite[i].ptr = 0;
            tapt1->subwhite[i].lower = 0;
            tapt1->subwhite[i].upper = 0;
        }
        free(struct whiteblock *tapt2->numwhite);
        free_ptr(tapt1);
    }
    /* print("after free white next = %d\n", tapt2); */
    /* print("after release white\n"); */
    free_rectblock(firstblock);
}

/* ..... */
/* ..... */
void free_section(firstsection)
struct sectionblock *firstsection
{
    struct sectionblock *tapt1, *tapt2;
    struct lineblock *tapt1, *tapt2;
    struct block_rect *tapt1, *tapt2;

    /* print("get into free section\n"); */
    for (tapt1 = firstsection->firstsubsection, tapt1 != NULL; ) {
        tapt2 = tapt1->next;
        /* print("section = %d section = %d\n", tapt1, tapt1->index); */
        for (tapt1 = tapt1->firstlineblock, tapt1 != NULL; ) {
            tapt2 = tapt1->next;
            /* print("tapt1 = %d line = %d first = %d current %d\n", tapt1,
                tapt1->index, tapt1->first_block->index,
                tapt1->current_block->index); */
            /*
                for (tapt1 = tapt1->first_block, tapt1 != NULL; ) {
                    tapt2 = tapt1->next;
                    free_block(tapt1);
                    tapt1 = tapt2;
                }
                tapt1->first_block = NULL;
            */
            /* print("free line tapt1 = %d tapt1->index = %d under %d t-cap %d\n",
                tapt1, tapt1->index, tapt1->under_section, tapt1->x_caption); */
            /* free(struct lineblock *tapt1); */
            /* print("after free line tapt2 = %d\n", tapt2); */
        }
    }
}

```

—262—

—262—

【図60】

```

        xs_set_foreground(TEXT_COLOR, wd);
        for (y = tmapsect->upper_y; y <= tmapsect->lower_y; y++)
            for (x = tmapsect->left_x; x <= tmapsect->right_x; x++)
                if ((tmapsect->image_ptr[x][y][tmapsect->text_color]) && (x > 0))
                    p[y][x] = TEXT_BLOCK;
                y1 = (int)((float)y/scale_displayform);
                x1 = (int)((float)x/scale_displayform);
                xs_picture(wd, x1, y1);
            }
        }

        /*..... get picture into picture display .....*/
        struct sectionblock *tmapsect2;
        void get_picture(tmapsect2, p, att, color, wd);
        struct sectionblock *tmapsect2;
        unsigned char *p;
        int att, color;
        widget wd;
        {
            int x, y, x1, y1;
            xs_set_foreground(TEXT_COLOR, wd);
            for (y = tmapsect2->upper_y; y <= tmapsect2->lower_y; y++)
                for (x = tmapsect2->left_x; x <= tmapsect2->right_x; x++)
                    if ((tmapsect2->image_ptr[x][y][tmapsect2->text_color]) && (x > 0))
                        p[y][x] = att;
                        y1 = (int)((float)y/scale_displayform);
                        x1 = (int)((float)x/scale_displayform);
                        xs_picture(wd, x1, y1);
                }
            }

        /*..... separate large .....*/
        void separate_large(firstsection, wd);
        struct sectionblock *firstsection;
        widget wd;
        {
            struct sectionblock *tmapsect, *tmapsect2, *tmapsect3;
            int x, y;
            for (tmapsect = firstsection->firstsection, tmapsect2 = NULL;
                get_text(tmapsect, h_image, wd, TEXT_COLOR);
                for (tmapsect2 = firstsection->firstsection; tmapsect2 != NULL;
                    tmapsect2 = tmapsect2->next;){
                    if (tmapsect2->text_color == TEXT_COLOR){
                        draw_line(tmapsect2, wd, LINE_COLOR, 1);
                    } else if (tmapsect2->text_color == LINE_COLOR, 3){
                        draw_line(tmapsect2, wd, LINE_COLOR, 3);
                    } else if (tmapsect2->text_color == LINE_COLOR, 2){
                        draw_line(tmapsect2, wd, LINE_COLOR, 2);
                    } else if (tmapsect2->text_color == LINE_COLOR, 4){
                        draw_line(tmapsect2, wd, LINE_COLOR, 4);
                    }
                }
            }

        /* print('get out line picture\n'); */
        /*..... get picture into picture display .....*/
        void get_line(tmapsect, p, att, color, wd);
        struct sectionblock *tmapsect;
        unsigned char *p;
        int att, color;
        widget wd;
        {
            int x, y, x1, y1;
            xs_set_foreground(TEXT_COLOR, wd);
            for (y = tmapsect->upper_y; y <= tmapsect->lower_y; y++)
                for (x = tmapsect->left_x; x <= tmapsect->right_x; x++)
                    if ((tmapsect->image_ptr[x][y][tmapsect->text_color]) && (x > 0))
                        p[y][x] = att;
                        y1 = (int)((float)y/scale_displayform);
                        x1 = (int)((float)x/scale_displayform);
                        xs_picture(wd, x1, y1);
                }
            }
    }

```

【图 6-1】

[illegible]

【図64】

```

)
/* allocate a double matrix with ranges (nrl..nrh, ncl..nch). */
double **matrix(nrl, nrh, ncl, nch)
{
    int nrl, nrh, ncl, nch;

    int i;
    double **m;
    m = (double **) malloc((nrh-nrl+1)*sizeof(double *));
    if (!m) perror("allocation failure 1 in matrix");
    m += nrl;
    for (i = nrl; i <= nrh; i++) {
        if (!m[i]) malloc((nch-ncl+1)*sizeof(double));
        if (!m[i]) perror("allocation failure 2 in matrix");
        m[i] += ncl;
    }
    return m;
}

/* deallocate a float vector */
void free_vector(f, nl, nh)
{
    float *v;
    int nl, nh;
    free((float *) (v+nl));
}

/* deallocate an int vector */
void free_vector_i, nl, nh)
{
    int *v;
    nl, nh;
    free((int *) (v+nl));
}

/* deallocate a double vector */
void free_vector(d, nl, nh)
{
    double *v;
    int nl, nh;
    free((double *) (v+nl));
}

/* deallocate a float matrix */
void free_matrix(f, nrl, nrh, ncl, nch)
{
    float **m;
    int nrl, nrh, ncl, nch;

    int i;
    for (i = nrl; i <= nrh; i++) free((float *) (m[i]+ncl));
    free((float *) (m+nrl));
}

/* deallocate an int matrix */
void free_matrix_i, nrl, nrh, ncl, nch)
{
    int **m;
    int nrl, nrh, ncl, nch;

    int i;
    for (i = nrl; i <= nrh; i++) free((int *) (m[i]+ncl));
    free((int *) (m+nrl));
}

```

```

/* deallocate an char matrix */
void free_matrix(c, nrl, nrh, ncl, nch)
{
    unsigned char **m;
    int nrl, nrh, ncl, nch;

    int i;
    for (i = nrl; i <= nrh; i++) free((char *) (m[i]+ncl));
    free((char *) (m+nrl));
}

/* deallocate a double matrix */
void free_matrix(d, nrl, nrh, ncl, nch)
{
    double **m;
    int nrl, nrh, ncl, nch;

    int i;
    for (i = nrl; i <= nrh; i++) free((double *) (m[i]+ncl));
    free((double *) (m+nrl));
}

```

-267-

-267-

【図66】

```

        struct data_outline *new_data_outline(int, int)
    {
        int n1, n2;
        struct data_outline *v;
        v = (struct data_outline *) malloc((unsigned)(n1+n2+1));
        if (!v) {
            printf("allocation error: data_outline\n");
            exit(1);
        }
        v->v = v;
        for (i = n1; i < n2; i++) {
            v[i].total_left = v[i].total_right = 0;
            v[i].current_left = v[i].current_right = NULL;
            v[i].headleft = v[i].headright = NULL;
        }
        return v;
    }

    ..... Free pair in block
    .....
    void free_pair(pair1, pair2, n1, n2)
    {
        int i;
        if (pair1 == NULL) return;
        for (i = n1; i < n2; i++) {
            free((int *)pair1[i]);
            free((int *)pair2[i]);
        }
        free((int **)pair1);
        free((int **)pair2);
        pair1 = pair2 = NULL;
    }

    ..... Free block_rect
    .....
    void free_block_rect(struct block_rect *block)
    {
        free(pair(block->pairright, block->pairleft, block->upper_y, block->lower_y,
            block->block_rect->nblock));
        block = NULL;
    }

    ..... Free white
    .....
    void free_white(struct white *white)
    {
        free(pair(white->pairright, white->pairleft, white->upper_y, white->lower_y,
            white->nblock));
        white = NULL;
    }

    ..... Allocate memory for data_outline
    .....
    ..... Adding one element to data_outline on right side
    .....

```


【図68】

```

void set_outline(x, y, nblook)
int x, y;
struct whiteblock *nblook;
{
    if (x < pairright[0]) {
        capair[numpair-2] = pairleft[0];
        capair[numpair-1] = pairleft[1];
        for (i = 0; i < numpair[0]; i++)
            pairright[i] = pairleft[i];
        for (i = 1; i < numpair[0]; i++)
            capair[numpair-2+i] = pairleft[i];
        else
            capair[numpair-2+i] = pairleft[pairleft[0]];
        } else {
            if (pairright[0] < pairleft[0])
                capair[numpair-2] = pairright[0];
                capair[numpair-1] = pairright[1];
                break;
            } else {
                pairright[0] = pairleft[0];
                pairright[1] = pairleft[1];
                pairleft[0] = capair[0];
                pairleft[1] = capair[1];
            }
        }
    }
    pairright[0] = pairleft[0];
    pairright[1] = pairleft[1];
    pairleft[0] = capair[0];
    pairleft[1] = capair[1];
    free((line *)capair);
}

/* Search the direction of next 'core'
..... */
void rightmost_core(p, old_direct, i, j, x1, y1, x2, y2, new_direct, newx, newy,
location)
{
    unsigned char *p;
    int i, j, x1, y1, x2, y2;
    int new_direct, newx, newy;
    int recording;
    int l, h;
    for (l = 1; l < search_map[old_direct][0]; l++) {
        new_direct = search_map[old_direct][l];
        newx = l * search_center[old_direct] * x;
        newy = j * search_center[old_direct] * y;
        if ((p[search_map[old_direct][l]] < search_map[old_direct][0]) ||
            (newx > x1 || newx < x2 || newy > y1 || newy < y2))
            return;
    }
    newx = x1;
    newy = y1;
    location = x1;
}

/* Set up the extreme point for the searched white block
..... */

```


【图 6-9】

[illegible]

【図70】

```

        wblock->previous->next = wblock->next;
        wblock->next->previous = wblock->previous;
    }

    /*..... Remove task block .....*/
    void remove_task_block(struct block_rect *first, struct block_rect *head)
    {
        if (head->firstsubblock == first) {
            head->firstsubblock = first->next;
            if (first->next != NULL) first->next->previous = NULL;
        } else if (head->currentsubblock == first) {
            head->currentsubblock = NULL;
        } else if (head->currentsubblock == first) {
            first->previous->next = first->next;
            first->next->previous = first->previous;
        } else {
            first->previous->next = first->next;
            first->next->previous = first->previous;
        }
    }

    /*..... Remove nontask block .....*/
    void remove_nontask_block(struct block_rect *first, struct block_rect *head)
    {
        if (head->firstsubblock == first) {
            head->firstsubblock = first->next;
            if (first->next != NULL) first->next->previous = NULL;
        } else if (head->currentsubblock == first) {
            head->currentsubblock = first->next;
            head->currentsubblock->previous = first->previous;
        } else {
            first->previous->next = first->next;
            first->next->previous = first->previous;
        }
    }

    /*..... Transfer block from one head to another head block .....*/
    void transfer_block(struct block_rect *front, struct block_rect *total, struct block_rect *first, struct block_rect *head)
    {
        /*print("get into transfer block\n");*/
        if (front->next != NULL) {
            remove_task_block(front, originalhead);
            remove_nontask_block(front, originalhead);
        }
        add_task_block(total, first, head);
        if (originalhead->next != NULL) {
            originalhead->next->previous = originalhead->next->previous;
        }
    }
}

```


—274—

[illegible]

—275—

```

}

//..... search black in white
//.....

int block_in_white(nblock, p, tapwhite)
struct block * nblock;
uint8_t (*tapwhite);
{
    int i, j, k, getin, block_no, keep;
    struct block * cur_block;

    printf("get into block_in_white\n");
    for (i = tapwhite->upper_y; i <= tapwhite->lower_y; i++) {
        printf("%d \n", i);
        for (j = 0; j <= tapwhite->pairright[i][0]; j++)
            printf("%d ", tapwhite->pairright[i][j]);
        printf("\n");

        for (k = j; k <= tapwhite->pairleft[i][0]; k--)
            printf("%d ", tapwhite->pairleft[i][k]);
        printf("\n");

        set_k(tapwhite->pairright[i], tapwhite->pairleft[i])

        printf("pair = %d \n", tapwhite->pairright[i][0]);
        printf("left = %d \n", tapwhite->pairleft[i][0]);
        printf(".....\n");
        for (j = j; j <= tapwhite->pairright[i][0]; j++)
            printf("%d ", tapwhite->pairright[i][j]);
        printf("\n");
    }
}

for (i = tapwhite->upper_y; block_no = 0; i <= tapwhite->lower_y; i++) {
    for (j = j; j <= tapwhite->pairright[i][0]; j++)
        printf("%d left = %d right = %d \n", i, tapwhite->pairright[i][0], j);

    for (k = (tapwhite->pairright[i][j]-1), getin = 0;
        k <= tapwhite->pairleft[i][j]; k++) {
        if ((getin == 0) && (j[i][k] != WHITE_BLOCK)) {
            getin++;
        } else if ((getin == 1) && (j[i][k] == LEFT_BLOCK)) {
            getin++;
        } else if ((getin == 0) && (j[i][k] == ONES)) {
            cur_block = new_block();
            cur_block->nblock = tapwhite->left_x;
            cur_block->p = tapwhite->left_x;
            tapwhite->upper_y = tapwhite->right_x;
            tapwhite->lower_y = cur_block->k;

            printf("after search_on_u")
        }

        printf("index = %d in_group = %d", cur_block->index, cur_block->in_group);
        for (l = (tapwhite->pairleft[i][j]+1) &&
            (j[i][l] != ONES)); l++;)
            return white;
}

```

—276—

[illegible]

[圖 7 5]

```

..... count continuous ones and zero
.....
/* print("out of block in range\n"); */

..... count the density inside the block
.....
void sort_block_rect(nblock, p, avg_one, std_one, avg_zero, std_zero)
.....
{
    struct block_rect nblock;
    unsigned char *p;
    float *avg_one, *std_one, *avg_zero, *std_zero;

    int i, j, k;
    int one_count, zero_count;
    int one, zero;
    int no_one, no_zero;

    /* print("place 3\n"); */

    one = i * vector(0, nblock->length * nblock->width - 1);
    zero = i * vector(0, nblock->length * nblock->width - 1);
    no_one = 0;
    no_zero = 0;

    for (j = nblock->upper_y; j <= nblock->lower_y; j++) {
        one_count = 0;
        zero_count = 0;
        for (k = 1; k <= nblock->pairing[1](0); k++) {
            for (l = nblock->pairing[1](1); l <= nblock->pairing[1](2); l++) {
                if (p[l] != 0) {
                    one_count++;
                    if (zero_count != 0) {
                        zero(no_one) = zero_count;
                        zero_count = 0;
                    }
                } else {
                    zero_count++;
                    if (one_count != 0) {
                        one(no_one) = one_count;
                        one_count = 0;
                    }
                }
            }
        }
        if (one_count != 0) {
            one(no_one) = one_count;
            one_count = 0;
        }
        if (zero_count != 0) {
            zero(no_zero) = zero_count;
            zero_count = 0;
        }
    }

    /* print("no_one = %d no_zero = %d\n", no_one, no_zero); */

    avg_one = avg_zero = std_one = std_zero = 0;
    if (no_one > 0) avg_one = one(no_one) / no_one;
    if (no_zero > 0) avg_zero = zero(no_zero) / no_zero;
    std_one = std_zero = 0;
    print("avg_one = %f avg_zero = %f\n", avg_one, avg_zero);
    /* print("std_one = %f std_zero = %f\n", std_one, std_zero); */

    free_vector(one, 0, nblock->length * nblock->width - 1);
    free_vector(zero, 0, nblock->length * nblock->width - 1);

    ..... count continuous ones and zero
    .....
}

```


[illegible]

—280—

【图 80】

[illegible]

—284—

【图 8-3】

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

—286—

[illegible]

【图 8 5】

[illegible]

【图 8-7】

[illegible]

【图 89】

—290—

【圖 90】

[illegible]

【图 9-1】

[illegible]

【図92】

```

return 0;
if (tapline->supery > max((firstline->lowery, secondline->lowery)
    < (lastline->lowery)))
    break;
else if ((!(overlap(tapline->left_x, tapline->right_x,
    firstline->left_x, firstline->right_x)) ||
    !(overlap(tapline->left_x, tapline->right_x, secondline->left_x,
    secondline->right_x))))
    if (neighbor == 1)
        break;
    }
}

if (if_neighbor == 1)
    print("bad add, ad tapad sized y2-ad 1 = %d\n",
        firstline->index, firstline->supery, firstline->lowery,
        secondline->index, secondline->supery, secondline->lowery,
        tapline->index, tapline->supery,
        tapline->left_x, tapline->right_x, tapline->lowery, tapline->index);
    if (neighbor == 0)
        if (if_neighbor == 1)
            break;
        }
    }

if (if_neighbor == 1)
    if (if_neighbor == 1)
        (cover_func(tapline->supery, firstline->lowery,
            secondline->lowery, secondline->supery)) == LINDOVELLAP))
            (cover_func(tapline->supery, firstline->lowery,
                secondline->lowery, secondline->supery)) == LINDOVELLAP))
            (min((firstline->width, secondline->width) - tapline->width)) <
            print("firstline = %d, secondline = %d, tapline = %d\n", firstline->index,
                secondline->index, tapline->index);
            return 1;
        }
    }
    return 0;
}

/* sort line
.....
struct lineblock *firstlineblock;
void sort_line(firstlineblock)
{
    struct lineblock *tap;
    int i, num, j, stop;
    for (num = 0, tap = firstlineblock->firstsubline, tap != NULL, tap = tap->next;
        num++;
    if (num > 0) {
        line_x = new_linepointer(0, num-1);
        for (i = 0, tap = firstlineblock->firstsubline, tap != NULL, tap = tap->next;
            i < num; i++)
            line_x[i] = tap;
        /* extra testing */
        sort_linepointer(line_x, num);
        firstlineblock->firstsubline = line_x[0];
        firstlineblock->currentsubline = line_x[num-1];
        if (num > 1)

```


—295—

[illegible]

—296—

[illegible]

【図96】

```

                                tapline=tapline->next){
    for ( i = 0; i <= (level-1)*8; i++){
        fprintf(line_file, " ");
        fprintf(line_file, "%d nontext index = %d ", tapline, tapline->index);
        fprintf(line_file,
            "x%d y%d l%d w%d att = %x block index%d index2%d g = %d l = %d r = %d\n",
            tapline->left_x, tapline->upper_y, tapline->length,
            tapline->width, tapline->att, tapline->first_block->index,
            tapline->current_block->index, tapline->group, tapline->l_caption,
            tapline->r_caption);
        print_ln(nextlevel, tapline);
    }

    /*----- Open the file for the output of line message -----*/
    /*----- Print the line information into a file -----*/

void print_lineblock(firstline)
struct lineblock *firstline;
{
    line_file = fopen("line_file.dat", "w");
    print_ln(1, firstline);
    fclose(line_file);
}

```

【図118】

```

                                tapsection = tapsection->next){
    if ((tapsection->att & (TITLE|NSOLID|VSOLID|PICTURE)) continue;
    else
        ini_section_block(tapsection, tapsection->firstlineblock, p,
                           column);

    /*----- Print the line information into a file -----*/

void print_sec(level, firstsec)
int level;
struct sectionblock *firstsec;
{
    int nextlevel, i;
    struct sectionblock *tapsubsec, *tapsec;
    nextlevel = level+1;
    for (tapsubsec = firstsec->firstsubsec; tapsubsec != NULL;
        tapsubsec = tapsubsec->next){
        for ( i = 0; i <= (level-1)*8; i++){
            printf(" ");
            printf("index = %d ", tapsubsec->index);
            printf("x%d y%d l%d w%d att = %x\n",
                tapsubsec->left_x, tapsubsec->upper_y,
                tapsubsec->length, tapsubsec->width, tapsubsec->att);
        }

        for (tapsec = firstsec->firstnontextsection; tapsec != NULL;
            tapsec = tapsec->next){
            for ( i = 0; i <= (level-1)*8; i++){
                printf(" ");
                printf("nontext index = %d ", tapsec->index);
                printf("x%d y%d l%d w%d att = %x\n",
                    tapsec->left_x, tapsec->upper_y, tapsec->length,
                    tapsec->width, tapsec->att);
                print_sec(nextlevel, tapsec);
            }
        }

        /*----- Open the file for the output of line message -----*/
        /*----- Print the section information into a file -----*/

void print_secblock(firstsec)
struct sectionblock *firstsec;
{
    print_sec(1, firstsec);
}

```

【図217】

Source Code for Mono-Spaced ("Courier") Segmentation

【图 9-7】

```

for ( i = n1, l = 0; l < n2; l++)
    v[i] column_no = 0;
return v;
}

..... allocation sectionblockpointer
.....

struct sectionblock **new_sectionpointer(n1, n2)
{
    int n1, n2;

    struct sectionblock **v;

    v = (struct sectionblock **) malloc((n2-n1+1)*sizeof(struct sectionblock));
    if (!v)
        printf("error: in the allocation of sectionpointer\n");
    exit(1);

    v = v+n1;

    return v;
}

..... sort line block pointer array
.....

void sort_linepointer(line, num)
struct lineblock **line;
int num;
{
    struct lineblock *tmp;

    for ( i = 0; i < num-1; i++) {
        for ( j = i+1; j < num; j++) {
            if (line[j]-supper_y > line[i]-supper_y) {
                tmp = line[i];
                line[i] = line[j];
                line[j] = tmp;
            }
        }
        if (i % 10 == 0) break;
    }
}

..... line block sorting
.....

struct lineblock **sort_y(firstlineblock, numline)
int *numline;
{
    struct lineblock *tmp;
    struct lineblock **line_x;
    int i, j, swap;

    *numline = 0;
    for ( tmp = firstlineblock->firstsubline; tmp != NULL; tmp = tmp->next)
        (*numline)++;
    for ( numline = 0; numline < (*numline); numline++)
        line_x = new_linepointer(0, (*numline)-1);
}

```

【図98】

```

for (i = 0; cap = firstlineblock->firstsubline; cap != NULL; cap = cap->next)
    line_x[cap] = cap;
/* extra testing */
if ((firstlineblock->currentsubline != NULL) &&
    (i != firstlineblock->currentsubline->index))
    printf("error in sort_line");
for (cap = firstlineblock->firstsubline; cap != NULL; cap = cap->next)
    line_x[cap] = cap;
sort_linepointer(line_x, "mainline");
return line_x;
}
/*..... section block setting
/*.....*/
struct sectionblock *sort_section(first, sum)
int *sum;
{
    struct sectionblock *cap;
    struct sectionblock *line_x;
    int i, j, stop;
    *sum = 0;
    for (cap = first->firstsubsection; cap != NULL; cap = cap->next)
        for (line_x[cap] = cap;
            for (i = 0; i < (*sum) - 1; i++) {
                for (j = 0; stop = 0; j = (*sum) - i - 1; j++)
                    if ((line_x[j]->upper_j > line_x[j+1]->upper_j) ||
                        (line_x[j]->lower_j > line_x[j+1]->lower_j)) {
                        cap = line_x[j+1];
                        line_x[j+1] = line_x[j];
                        line_x[j] = cap;
                        stop = 1;
                    }
                if (stop == 0) break;
            }
        return line_x;
    }
/*..... allocate new line block
/*.....*/
struct sectionblock *new_section(int, nh)
int nh, nh;
{
    struct sectionblock *v;
    v = (struct sectionblock *) malloc(sizeof(struct sectionblock));
    if (!v)
        printf("allocation error: section block\n");
    exit(1);
}

```

```

}
for (i = 0; i < cap; i++) {
    v[i].firstlineblock = v[i].currentlineblock = NULL;
    v[i].firstsubsection = v[i].currentsubsection = NULL;
    v[i].firstsubline = v[i].currentsubline = NULL;
    v[i].previous = v[i].next = v[i].left_x = v[i].right_x = 0;
    v[i].width = v[i].length = 0;
    v[i].att = UNKNOWN;
    v[i].mainlineblock = 0;
    v[i].index = 0;
    v[i].boundary_index = -1;
}
return v;
}
/*..... Chain the context section block
/*.....*/
struct sectionblock *new_section(headsection)
struct sectionblock *headsection;
{
    struct sectionblock *n;
    n = new_sectionblock(0, 0);
    if (headsection->firstsubsection == NULL)
        headsection->firstsubsection = n;
    else
        headsection->currentsubsection->next = n;
    n->previous = headsection->currentsubsection;
    headsection->currentsubsection = n;
    return n;
}
/*..... Chain the section block
/*.....*/
struct sectionblock *new_section(headsection)
struct sectionblock *headsection;
{
    struct sectionblock *n;
    n = new_sectionblock(0, 0);
    if (headsection->firstsubsection == NULL)
        headsection->firstsubsection = n;
    else
        headsection->currentsubsection->next = n;
    n->previous = headsection->currentsubsection;
    headsection->currentsubsection = n;
    return n;
}
/*..... take lineblock from section
/*.....*/
void take_section(line, currentsection)
struct lineblock *line;
struct sectionblock *currentsection;

```


【図100】

```

newhead->first_block->previous = line_tran->current_block;
line_tran->current_block->next = newhead->first_block;
newhead->first_block = line_tran->first_block;
newhead->next_block = line_tran->next_block;
if (line_tran->left_x < newhead->left_x)
    newhead->left_x = line_tran->left_x;
if (line_tran->right_x > newhead->right_x)
    newhead->right_x = line_tran->right_x;
if (line_tran->upper_y < newhead->upper_y)
    newhead->upper_y = line_tran->upper_y;
if (line_tran->lower_y > newhead->lower_y)
    newhead->lower_y = line_tran->lower_y;
newhead->width = newhead->right_x - newhead->left_x;
newhead->length = newhead->upper_y - newhead->lower_y;
}
//..... combination of line block and head
//.....
void bline_combine(newhead, line_tran)
struct lineblock *newhead, *line_tran;
{
    newhead->current_block->next = line_tran->first_block;
    line_tran->first_block->previous = newhead->current_block;
    newhead->current_block = line_tran->current_block;
    newhead->next_block = line_tran->next_block;
    if (line_tran->left_x < newhead->left_x)
        newhead->left_x = line_tran->left_x;
    if (line_tran->right_x > newhead->right_x)
        newhead->right_x = line_tran->right_x;
    if (line_tran->upper_y < newhead->upper_y)
        newhead->upper_y = line_tran->upper_y;
    if (line_tran->lower_y > newhead->lower_y)
        newhead->lower_y = line_tran->lower_y;
    newhead->width = newhead->right_x - newhead->left_x;
    newhead->length = newhead->upper_y - newhead->lower_y;
}
//..... line remove from some head and merge with another head
//.....
void line_remove(line_tran, originalhead, newhead, att)
struct lineblock *line_tran, *originalhead, *newhead;
int att;
{
    line_tran->originalhead = line_tran;
    bline_combine(newhead, line_tran);
    free(originalhead->line_tran);
}
//..... test if one line lies inside another line
//.....
int if_inside(firstline, secondline)
struct lineblock *firstline, *secondline;
{
    if (((float)cover_range(firstline->upper_y, firstline->lower_y,
        secondline->upper_y, secondline->lower_y) > 1.0) &&
        ((float)cover_range(firstline->left_x, firstline->right_x,
            secondline->left_x, secondline->right_x) > 1.0))
        return 1;
    return 0;
}
}
//..... cover part
//.....
float cover_part(mark, a1, a2)
int *mark;
int a1, a2;
{
    int unmark, i;
    for (i = a1; unmark = 0; i <= a2; i++)
        if (mark[i] == 0)
            unmark++;
        mark[i] = 1;
    return ((float)unmark/(float)(a2-a1+1));
}
//..... the number of "neighbor" on the left side
//.....
void test_left_neighbor(line, leftlabel, leftlength, head, line_list, nodistroy)
struct lineblock *line;
int *leftlength;
float *leftlabel;
struct lineblock **line_list;
struct lineblock **nodistroy;
{
    int n, upper, lower, mark, mostleft;
    struct nameline *aoline, *aliline, *tap;
    upper = line->current->upper_y;
    lower = line->current->lower_y;
    for (tap = line->previous; tap != NULL; tap = tap->previous)
        if (tap->current == NULL) continue;
        if (tap->current->upper_y < upper) upper = tap->current->upper_y;
        if (tap->current->lower_y > lower) lower = tap->current->lower_y;
    mark = (int *)vector(upper, lower);
    mostleft = line->current->left_x;
    aoline = aliline = NULL;
    leftlength = 0;
    for (tap = line->previous; tap != NULL; tap = tap->previous)
        if (tap->current == NULL) continue;
        if ((tap->current->right_x - mostleft) < leftlength)
            (cover_range(tap->current->upper_y, tap->current->lower_y,
                line->current->upper_y, line->current->lower_y) > 0.0)
            if ((n = cover_part(mark, tap->current->upper_y, tap->current->lower_y))
                > OVER_2)
            {
                if (tap->current->left_x < mostleft)
                    mostleft = tap->current->left_x;
                if (tap->current->right_x > leftlength)
                    leftlength = tap->current->right_x - mostleft;
                if (tap->current->upper_y < upper)
                    upper = tap->current->upper_y;
                if (tap->current->lower_y > lower)
                    lower = tap->current->lower_y;
                if (aoline == NULL)
                    aoline = tap;
                if (aliline == NULL)
                    aliline = tap;
            }
}

```


—303—

[illegible]

—304—

—304—

```

currentline = first;
top = line_y_sort(j)--upper_y;
bottom = line_y_sort(i)--lower_y;
for (j = i-1; j < num; j--) {
    if ((line_y_sort[j]) == NULL) continue;
    if ((lowerip(line_y_sort[j])--upper_y, line_y_sort(j))>lowery,
        upperip(bottom)--upper_y, top--bottom)) continue;
    if ((line_y_sort(j)--act == UNKNOW)) {
        currentline=bottom = new_base[line_y_sort(j)];
        currentline=length < MAXPOINT/SCALE_JAINTO;
        currentline=currentline+next;
        if (bottom < line_y_sort(j)--lowery)
            bottom = line_y_sort(j)--lowery;
    } else if ((line_y_sort(j)--upper_y > bottom) break;
}
for (j = i-1; j > 0; j--) {
    if ((line_y_sort[j]) == NULL) continue;
    if ((lowerip(line_y_sort(j)--upper_y, line_y_sort(j))>lowery,
        upperip(bottom)--upper_y, top--bottom)) continue;
    if ((line_y_sort(j)--act == UNKNOW)) {
        currentline=bottom = new_base[line_y_sort(j)];
        currentline=length < MAXPOINT/SCALE_JAINTO;
        currentline=currentline+next;
        if (bottom < line_y_sort(j)--lowery)
            bottom = line_y_sort(j)--lowery;
    } else {
        if ((line_y_sort(j)--lowery < top+1)
            break;
        sort_saseline(first);
        index_test(first, firstline, line_y_sort);
        for (currentline = first; currentline != NULL;
            currentline=currentline+next){
            if ((currentline==current==NULL) ||
                ((currentline==current==act)&&TEXT)) {
                currentline=current+act+MAXDOWN;
                test_neighbors(currentline, leftlevel, rightlength,
                    rightlength, firstline, line_y_sort, NULL);
                if (leftlevel <= rightlevel) & leftlength < rightlength & leftlength < rightlength & leftlevel <= rightlevel, rightlength, rightlength,
                    currentline=current+index, leftlevel, rightlevel, leftlength, rightlength, rightlength;
                if (leftlevel >= 2*(rightlength+1)) {
                    currentline=current+act + TEXT/PICTURE/RECTEST;
                }
                if (leftlevel >= 3) {
                    currentline=current+act + PICTURE/ADDTEXT/TEXT;
                } else {
                    lefttitle=currentline, line_y_sort, firstline;
                    test_right_neighbors(currentline, rightlength,
                        rightlength, firstline, line_y_sort, NULL);
                }
                if (rightlevel >= 1) {

```

[illegible]

—308—

[illegible]

【圖 1 0 9】

```

if (yoverlap == 1.0) {
    if (align == TOPLEFT) {
        x1 = bottomsec->left_x;
        x2 = topsec->left_x;
        y1 = topsec->bottom_y;
        y2 = bottomsec->top_y;
    } else {
        x1 = topsec->right_x;
        x2 = bottomsec->right_x;
        y1 = topsec->bottom_y;
        y2 = bottomsec->top_y;
    }
} else if (align == TOPRIGHT) {
    x1 = topsec->right_x;
    x2 = bottomsec->right_x;
    y1 = topsec->bottom_y;
    y2 = bottomsec->top_y;
} else {
    x1 = topsec->left_x;
    x2 = bottomsec->left_x;
    y1 = topsec->bottom_y;
    y2 = bottomsec->top_y;
}

count = 0;
for (s = x1; s <= x2; s++) {
    for (y = y1; y <= y2; y++) {
        if ((original[s][y] & 0x01) == 0x01) {
            count++;
        }
    }
}

if (count > 0) return 1;
if (count > 0) return 1;

printf("place 2 x1 = %d x2 = %d y1 = %d y2 = %d\n", x1, x2, y1, y2);

/*
if (x2 > x1)
for (tmp = first->firstsubsec; tmp != NULL; tmp = tmp->next) {
    if ((tmp->bottomsec) != (tmp->topsec)) continue;
    topcover = cover_range(tmp->left_x, tmp->right_x, y1, y2);
    bottomcover = cover_range(tmp->left_x, tmp->right_x, x1, x2);
    if (((topcover == 0.0) || (topcover == 1.0)) ||
        ((bottomcover == 0.0) || (bottomcover == 1.0))) continue;
    else return 1;
}

/* printf("x1 = %d x2 = %d y1 = %d y2 = %d\n", x1, x2, y1, y2); */

/*
for (tmp = first->firstsubsec; tmp != NULL; tmp = tmp->next) {
    topcover = cover_range(tmp->left_x, tmp->right_x,
        tmp->left_x, tmp->right_x);
    bottomcover = cover_range(tmp->left_x, tmp->right_x,
        tmp->left_x, tmp->right_x);
    if (((topcover == 0.0) || (topcover == 1.0)) ||
        ((bottomcover == 0.0) || (bottomcover == 1.0)))
        continue;
    else return 1;
}

/*
if ((tmp->lower_y == topsec->upper_y) ||
    (tmp->upper_y == bottomsec->lower_y) ||
    (tmp->left_x == max(topsec->right_x, bottomsec->right_x)))
    return 1;
}

/*

```


【图 1 1 2】

[illegible]

—314—

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```


【例 115】

[illegible]

[図120]

```

for (tap = white, num = 0, top = NULL; tap = tap->next, num++;
    tap[num] = tap;
    for (i = 0, j = num; i++;
        for (k = 0, stop = 0; k < num - j - 1; k++)
            if (tap[j+k]-white-seed_right > tap[j]-white-seed_right) {
                tap = tap[j+k];
                tap[j] = tap;
                stop = 1;
            }
        if (stop == 0) break;
        top3 = -1;
        for (i = 0, j = num; i++;
            if (tap[i]-white-seed_right < tap[j]-white-seed_right+OFFSET) {
                if (tap[i]-white-seed_right > top3)
                    top3 = tap[i]-white-seed_right;
            }
            else {
                top2 = tap[i]-white-seed_right;
                break;
            }
        }
        if (top3 == -1) top3 = top1;
    }
}

.....
.....
void first_y_block(white, top1, top3)
{
    struct white_for_table *white;
    int num, stop, i, j;

    for (tap = white, num = 0, top1 = NULL, top = tap->next, num++;
        tap = (struct white_for_table *)first_y_block(tap->next, num);
        for (tap = white, num = 0; tap != NULL; tap = tap->next, num++;
            for (i = 0, j = num; i++;
                for (k = 0, stop = 0; k < num - j - 1; k++)
                    if (tap[j+k]-white-seed_lowery > tap[j]-white-seed_lowery) {
                        tap = tap[j+k];
                        tap[j] = tap;
                        stop = 1;
                    }
                if (stop == 0) break;
                top3 = -1;
                for (i = 0, j = num; i++;
                    if (tap[i]-white-seed_lowery < tap[j]-white-seed_lowery+OFFSET) {
                        if (tap[i]-white-seed_lowery > top3)
                            top3 = tap[i]-white-seed_lowery;
                    }
                    else {
                        top2 = tap[i]-white-seed_lowery;
                        break;
                    }
                }
                if (top3 == -1) top3 = top1;
            }
        }
    }
}

```


【図125】

```

/*****
***** File: block_util.c
***** Author: Shin-Yuan Wang
***** Date : 1-28-92
***** Copyright 1992 Canon Information Systems
*****/

/***** test if covered range overlap *****/

int if_overlap(a1, a2, b1, b2)
{
    int a1, a2, b1, b2;
    if ((a2 < b1) || (b2 < a1)) return 0;
    return 1;
}

/***** minimum of two integers *****/

int min(a, b)
{
    int a, b;
    if (a <= b) return a;
    else return b;
}

/***** ratio of overlappint *****/

float cover_range(a1, a2, b1, b2)
{
    int a1, a2, b1, b2;
    float range = 0.0;
    int x1, x2;
    if ((a2 < b1) || (b2 < a1)) return range;
    else {
        x1 = max(a1, b1);
        x2 = min(a2, b2);
    }
    range = (float)(x2-x1-1)/(float)(a2-a1+1);
    return range;
}

```

【図131】

```

/*****
***** Filename: blockline.h
***** Author: Shin-Yuan Wang
***** Date: 1-15-92
***** Copyright 1992 Canon Information Systems
*****/

#define EMPTY 0
#define OVERLAP 1
#define NEXT 2
#define SURROUNDIN 3
#define SURROUNDOUT 4

#define TH 1.7
#define WIDTHSTD 1.2
#define CAPLENGTH 0

#define DOT_BOTTOM 0.3
#define DOT_WIDTH 2
#define LINEOVERLAP 0.5

#define LEFTGAP 0x00
#define RIGHTGAP 0x01
#define NOISE_NO 3

typedef int fillcondition;

extern void as_draw_box();
extern void convert_to_headblock();
extern void transfer_block();
extern void remove_textblock();
extern int max();
extern int min();
extern float cover_range();
extern struct blockline **new_linepointer();
extern void sort_linepointer();

```

【図127】

```

/*****
***** Filename: block_main.h
***** Author: Shin-Yuan Wang
***** Date: 1-11-92
***** Copyright 1992 Canon Information Systems
*****
*****/

#define TESTTIME 20
#define TEXT_BLOCK 0x11,
#define LINEPICTURE_BLOCK 0x21
#define MPICTURE_BLOCK 0x31
#define TABLE_BLOCK 0x41
#define LINE_BLOCK 0x51
#define FRAME_BLOCK 0x61
#define UNKNOWN_BLOCK 0x71

#define TEXT_COLOR 0
#define TABLETEXT_COLOR 1
#define LINE_COLOR 2
#define PICTURE_COLOR 3
#define MPICTURE_COLOR 4
#define UNKNOWN_COLOR 5
#define TABLE_COLOR 6
#define FRAME_COLOR 7
#define LINECAP_COLOR 8

enum {SP_READ = 0x01,
      SP_REDUCE = 0x02,
      SP_COPY = 0x04,
      SEP_BLOCK = 0x08};

struct block_image
{
    int width;
    int height;
    unsigned char** pixel;
};

int xoff[TESTTIME][2] = {(0, 1), (1, 2), (0, 1), (1, 3), (0, 1),
                          (1, 2), (0, 0), (0, 0), (1, 2), (1, 2),
                          (2, 2), (2, 2), (1, 0), (2, 1), (1, 0),
                          (3, 1), (0, 1), (0, 1), (1, 3), (1, 2)};

int yoff[TESTTIME][2] = {(0, 0), (0, 0), (1, 1), (1, 1), (2, 2),
                          (2, 2), (0, 1), (1, 2), (0, 1), (1, 2),
                          (0, 1), (0, 2), (0, 1), (0, 1), (1, 2),
                          (1, 2), (1, 2), (0, 1), (1, 2), (0, 1)};

#define TOTALCOLOR 7

char *color_label[] = {
    "Text",
    "Text in the Table",
    "Picture",
    "Line Drawing Graphics",
    "Solid Line (Vertical or Horizontal)",
    "Table",
    "Frame"
};

int color_code[] = {TEXT_COLOR, TABLETEXT_COLOR, MPICTURE_COLOR, PICTURE_COLOR,
                    LINE_COLOR, TABLE_COLOR, FRAME_COLOR};

```

【図132】

```

/*****
***** Filename: blocksection.h
***** Author: Shin-Yuan Wang
***** Date: 1-11-92
***** Copyright 1992 Canon Information Systems
*****
*****/

#define GAP 15
#define OVERLAP 0.8
#define NOVERLAP 0.5
#define LEVELOVERLAP 0.0
#define LEFTSIDE 0x10
#define LEFTALIGN 0x20
#define LEFTTOWLONGER 0x40
#define LEFTBOTLONGER 0x20
#define RIGHTSIDE 0x01
#define RIGHTALIGN 0x08
#define RIGHTTOWLONGER 0x04
#define RIGHTBOTLONGER 0x01
#define TEXTPIXEL 0x08

#define NOISE_COUNT 5

extern int line_inside();
extern int max();
extern int min();
extern int if_overlap();
extern void xa_draw_box();
extern float cover_range();
void exit();

struct segment_gap
{
    int column_no;
    int *begin_column;
    int *end_column;
};

```


【☒ 1 2 8】

—327—

【図135】

```

/*****
 * Filename: Table.h
 * Author: Shin-Yuan Wang
 * Date: 1-15-93
 * Copyright 1992 Canon Information Systems
 *****/

#include "blockgeneral.h"

#define NONE 0
#define OFFSET 3
#define OFFSET2 6
#define MEM 1
#define OLD 2
#define MONSTOP 0
#define STOP 1
#define TABLE_TESTTIME 2

struct white_for_table
{
    int status;
    struct whiteblock *white;
    struct white_for_table *previous, *next;
};

```

【図136】

```

/*****
 * Filename: images.h
 * Author: Shin-Yuan Wang
 * Date: 1-15-93
 * Copyright 1992 Canon Information Systems
 *****/

#include<stdio.h>
#include<stdlib.h>

struct images
{
    int sizeR;
    int sizeC;
    int dpi;
    unsigned char **pixel;
};

```

【図142】

```

/*
 * Filename: autoinput.h
 * Header file for autoinput.c
 */

/* Function Definitions for mymalloc.c */
#ifdef CPP
/* Definition for C++ files */
extern "C" {
    void auto_initialize();
    int auto_input_int();
    void auto_terminate();
}
#else
/* Definition for C files */
void auto_initialize();
int auto_input_int();
void auto_terminate();
#endif

```

【図145】

```

/*-----*/
void Zerovpp_terminate(mzlist)
    struct NonZerovppList *mzlist; /* The list */
{
    free_vector(mzlist->list); /* Deallocate memory */
    mzlist->allocated = 0; /* Set so won't be accidently used */
    mzlist->size = 0;
}

/*-----*/
void Zerovpp_add(mzlist, a, b)
    struct NonZerovppList *mzlist; /* The list */
    int a, b; /* Start and stop of the zerovpp */
{
    if (mzlist->size+2 > mzlist->allocated) {
        printf("NonZerovppList FULL!\n");
        exit(1);
    }
    mzlist->list[mzlist->size++] = a;
    mzlist->list[mzlist->size++] = b;
}

```

【図214】

```

/* If your routines are in C++ then CPP must be defined in make file */
#ifdef CPP
extern "C" void xa_set_foreground(int x, int y, int b);
extern "C" void xa_set_background(int x, int y, int b);
extern "C" void xa_init();
extern "C" void xa_flush();
extern "C" void xa_plotxy(int x, int y);
extern "C" void xa_xplotxy(int x, int y);
extern "C" void xa_redisplay();
extern "C" void xa_dia_bitmap(int x, int y, int width, int height, char *buf);
extern "C" void xa_clr_vin();
extern "C" void xa_draw_box(int x, int y, int width, int height);
extern "C" void xa_string(int x, int y, char *str);
extern "C" void xa_setcolor(int color);
extern "C" void xa_vin_dia(int width, int height);
#endif

```

【圖 138】

[illegible]

【図140】

```

    temp2 = *(oc3++);
    *(inc++) = (temp1 & *(oc1++)) | (temp2 & *(oc1++)) | (temp3 & *(oc3++));
    if (b) // Do the right end
        *(inc) = (*(oc1) | *(oc2) | *(oc3));
    }
    // Do the bottom line (if a == 1 or 2)
    nc = "nc";
    if (a == 1) // a == 1 it's one row
    {
        oc1 = "or";
        for (j=0; j<newc; j++)
        {
            temp1 = *(oc1++);
            *(inc++) = temp1 & *(oc1++);
            if (b) // Do the right end
                *(inc) = *(oc1);
            } else {
                if (a == 2) // a == 2 it's two rows
                {
                    oc1 = *(oc1++);
                    oc2 = *(oc2++);
                    for (j=0; j<newc; j++)
                    {
                        temp1 = *(oc1++);
                        temp2 = *(oc2++);
                        *(inc++) = (temp1 & *(oc1++)) | (temp2 & *(oc1++));
                        if (b) // Do the right end
                            *(inc) = *(oc1) | *(oc2);
                    }
                }
            }
        }
    }
}

oc2 = *(oc2++);
nc = "nc";
for (j=0; j<newc; j++)
{
    temp1 = *(oc1++);
    temp2 = *(oc2++);
    *(inc++) = (temp1 & *(oc1++)) | (temp2 & *(oc1++));
    if (b) // Do the right end
        *(inc) = (*(oc1) | *(oc2));
}
// Do the bottom line (if a == 1)
// a == 1 it's one row
if (a == 1)
{
    oc1 = "or";
    for (j=0; j<newc; j++)
    {
        temp1 = *(oc1++);
        *(inc++) = temp1 & *(oc1++);
    }
    if (b) // Do the right end
        *(inc) = *(oc1);
    }
}

// *****
// 1st AND in 1st
void find_and_1st(cimages oldline, cimages newline)
{
    int oldc, oldr; // size of old image
    int newc, newr; // size of new image
    int i, j, a, b;
    unsigned char *oc1, *oc2, *oc3, *nc; // The old & new column pointers
    unsigned char *inc; // The old & new row pointers
    register unsigned char temp1, temp2, temp3;

    oldc = oldline.size;
    oldr = oldline.alter; // Should be ==
    newc = newline.size;
    newr = (oldr-1)/2; // Should be ==
    newline.reallocate(newc, newr-1);

    a = oldc - newc; // How much is partially covered in the last row
    if (a == 1) // It's evenly aligned
        b = oldc - newc; // How much is partially covered in the last column
    if (b == 2)
    {
        newc++;
        b = 0; // == 0 for optimization of ends below
    }
    or = oldline.pixel;
    nr = newline.pixel;
    for (i=0; i<newr; i++) // Do the middle
    {
        oc1 = *(oc1++);
        oc2 = *(oc2++);
        oc3 = *(oc3++);
        nc = *(nc++);
        for (j=0; j<newc; j++)
        {
            temp1 = *(oc1++);
            temp2 = *(oc2++);

```

【図141】

```

/* filename: autolinput.h
 * This file contains the code to input from a filename
 * .....
#include <stdio.h>
#include "autolinput.h"

/* Global variables */
int auto_eof_detected;
FILE *auto_fp;
int save_count;
char filename[100];

void auto_initialize()
{
    int not_exit = 1;
    auto_eof_detected = 0;
    while(not_exit)
    {
        printf("Input filename for auto-input? ");
        scanf("%s", filename);
        auto_fp = fopen(filename, "r");
        if (auto_fp == NULL) /* Read open failed */
        {
            auto_fp = fopen(filename, "w");
            if (auto_fp == NULL) /* Write open failed */
            {
                printf("cannot open file %s\n", filename);
            }
            else /* Opened new file */
            {
                printf("Enter 1 at the next prompt.\n");
                not_exit = 1;
                auto_eof_detected = 1;
            }
        }
        else /* Read open succeeded */
        {
            printf("Opening existing file %s\n", filename);
            not_exit = 0;
        }
    }

    printf("Do you wish to save the auto-input into this file? (1 = yes) ");
    save_to_file = auto_input_int();
    printf("Manual Change the 1st '1' to '0' to turn off appending to the file\n");
    save_count = 0;
}

int auto_input_int()
{
    int i;
    if (auto_eof_detected)
    {
        printf("Manual Input 1-9 = newline, -9999 = exit) ? ");
        scanf("%d", &i);
        if (i == -9999)
        {
            fclose(auto_fp);
            exit(0);
        }
        if (save_to_file == 1)
        {
            if (i == -9) /* Print new lines */
            {
                printf(auto_fp, "\n"); /* Print the new lines */
                fflush(auto_fp);
                save_count = 0;
            }
            return auto_input_int(); /* And input again */
        }
        printf(auto_fp, "%d ", i);
        fflush(auto_fp);
        if (++save_count > 20)
        {
            save_count = 0;
            printf(auto_fp, "\n");
        }
        return i;
    }
    if (feof(auto_fp, "d", &i) == EOF)
    {
        printf("\n----- END OF FILE DETECTED -----");
        if (save_to_file == 1)
        {
            fclose(auto_fp);
            auto_fp = fopen(filename, "a");
            if (auto_fp == NULL)
            {
                printf("***** WARNING: CANNOT APPEND TO FILE. Is ....\n", filename);
                save_to_file == 0;
            }
            printf(auto_fp, "\n");
            auto_eof_detected = 1;
            return auto_input_int();
        }
        printf("\n\n");
        return i;
    }
}

printf("Do you wish to save the auto-input into this file? (1 = yes) ");
save_to_file = auto_input_int();
printf("Manual Change the 1st '1' to '0' to turn off appending to the file\n");
save_count = 0;
}

int auto_input_int()
{
    int i;
    if (auto_eof_detected)
    {
        printf("Manual Input 1-9 = newline, -9999 = exit) ? ");

```

【図143】

```

/* File name: charlist.c
 * This file contains Charlist and NonZeroHeight utility routines
 * .....
 */

/* .....
 * Routines to handle Character lists
 * .....
 */
/* By Christopher Sherrick */
#include "charlist.h"
#include "gmath.h"

/* .....
 * Initialize and allocate a CharcterList (set up curve pointers)
 * .....
 */
/* Call this routine once only! */
void CharListInit(CharList *clst, int new_vector_height)
{
    int clist_max; /* The CharacterList
    int vector_height; /* Height of the line (to allocate curves) */
    int i; /* A looping var */

    /* Allocate memory for characters */
    clist_max = (struct Character *) malloc((unsigned) clist_max
    * sizeof(struct Character));
    if (!clist_max) {
        fprintf(stderr, "Memory Allocation Error in CharListInit\n");
        exit(1);
    }
    clist_max = 0; /* Number of characters */
    clist_max_allocated = clist_max; /* Size Allocated */
    clist_max = 1; /* Set the first character */
    clist_max_vector_height = vector_height;

    /* Allocate memory for curves and set pointers */
    for (i = 0; i < clist_max; i++) {
        clist_max_vector_height = vector_height;
        clist_max_vector_height = vector_height;
    }
}

/* .....
 * Zero a previously allocated CharacterList */
void CharListZero(CharList *clst)
{
    struct Character *clst;
    clist_max = 0; /* Number of characters */
    clist_max_allocated = 0; /* No first character */
}

/* .....
 * Remove a character from the list
 * .....
 */
void CharListRemove(CharList *clst, int i)
{
    struct Character *clst;
    int i;
    for (i = clist_max; i > 0; i--) {
        if (i == i) {
            free_vector(clist_max_vector_height);
            free_vector(clist_max_vector_height);
        }
    }
}

```

```

}
free(struct Character *) clist_max;
clist_max = 0;
clist_max_allocated = 0;
clist_max_first = 0;
}

/* .....
 * Places a new character in the charlist:
 * .....
 */
/* Takes the character beyond the charlist (.chr[clst]).
 * calculates its size, and inserts it into the list */
void CharListInsert(CharList *clst, struct Character *clst)
{
    struct Character *newchar; /* Pointer to the new character adding */
    newchar = (struct Character *) malloc(sizeof(struct Character));
    /* Calculate the position of the new character */
    CharListRecalculateLoc(newchar);
    /* Check for full list */
    if (clist_max == clist_max_allocated) /* If too big. */
    {
        printf("Error in CharList_insert: Character list is FULL\n");
        exit(1);
    }
    /* Place new in the list character */
    CharListLocateCharacter(clst, newchar);
    /* This routine repositions a character in the charlist. It removes
    * the character, recalculates loc and inserts it in the list accordingly
    */
    void CharListRecalculateLoc(CharList *clst, struct Character *clst)
    {
        struct Character *clst;
        /* Recalculate the position of the character */
        CharListRecalculateLoc(clst);
        return;
    }
    /* Remove the character from the list */
    if (clst_max == 0) /* If list on list */
    {
        clist_max = clist_max_allocated;
        clist_max_first = 0;
    }
    else {
        clst_max_vector_height = clst_max;
        if (clst_max == 0)
            clst_max_vector_height = clst_max;
    }
    /* Place the character in the list */
    CharListLocateCharacter(clst, clst);
}

/* .....
 * Insert a character after doing a cut */

```


【図144】

```

/* This routine will insert the character beyond the charlist (chrlist).
 * loc is not used for both character and the new character, however
 * place the new character after chrlist. After finished, it is an
 * extremely good idea to resort the heaplist!
 */
void Charlist_insert_after_ptr(charlist, ch, after)
struct Character *chrlist; /* The character */
struct Character *after; /* Place the new character after this one */
{
    struct Character *newchar; /* Pointer to the new character adding */
    newchar = &(list->next); /* Set newchar */
    /* Calculate the position of the new character */
    Charlist_recalculate_loc(newchar);
    Charlist_recalculate_loc(ch, after);
    /* Check for full list */
    if (chrlist->next == NULL) /* Increment the size of the list */
    {
        printf("Error in Charlist_remove_char(). Character list is FULL!\n");
        exit(1);
    }
    newchar->prev = ch;
    newchar->next = ch->next;
    ch->next = newchar;
    if (newchar->next != 0)
        newchar->next->prev = newchar;
}

/* This routine will remove the character from the list.
 * Also, prev and next pointers of ch are NOT altered.
 */
void Charlist_remove_char(charlist, ch)
struct Character *chrlist;
struct Character *ch;
{
    /* Remove the character from the list */
    if (ch->prev == 0) {
        list->first = ch->next;
        ch->next->prev = 0;
    }
    else {
        ch->prev->next = ch->next;
        if (ch->next != 0)
            ch->next->prev = ch->prev;
    }
}

/* This routine will calculate the position of a character */
void Charlist_recalculate_loc(charlist, ch) /* Output is the position */
struct Character *ch; /* The character */
{
    int i, loc;
    /* Calculate the position of the new character */
    loc = 0;
    for (i = 0; i < list->size; i++) /* Position = average position... */
        loc += ch->next[i];
    loc /= list->size; /* Set it */
}

```

```

/* This routine will place a character (that is not in the charlist)
 * into the charlist.
 */
void Charlist_insert(charlist, ch)
struct Character *chrlist; /* The character */
struct Character *ch; /* The character */
{
    struct Character *p; /* Used for searching for location */
    struct Character *q; /* The character before p */
    /* Is the list empty? */
    if (list->size == 0)
    {
        ch->next = 0;
        ch->prev = 0;
        list->first = ch;
        return;
    }
    /* Possible optimization: Don't start from start of list...
     * start from where at */
    loc = ch->loc; /* Retrieve position */
    /* Is it at the beginning of the list? */
    p = list->first; /* Load p with the first item on the list */
    if (loc < p->loc) /* If it's the first on the list */
    {
        p->prev = ch;
        ch->next = list->first;
        list->first = ch;
        return;
    }
    /* Set the pointers */
    /* Set the loc on the list to this */
    while (p->loc <= loc) /* Until find record past where to place */
    {
        list->next = p; /* Remember last position */
        p = p->next; /* Get the next character */
        if (p == 0) /* If it's the end of list, insert it there */
            break;
    }
    ch->prev = list->next;
    ch->next = list->next->next;
    list->next->next = ch;
    if (p->loc == loc) /* If newchar isn't at the end of the list */
        p->prev = ch;
}

/* Allocate memory for list */
size = size; /* Compute memory allocated, etc */
list->next = vectorize(list); /* Allocate memory
 * for the list */
list->size = 0; /* Set to zero elements */

```

【图 146】

—336—

```

// Filename: climage.cop
// This file contains a vast array image utility functions, for a cimage
//
//
//
//
// underline this co make limage.c run invisibly (except for error)
//define PRINT_IMAGES_DEBUG

#include "climage.h"
#include "ia.h"

//new CLASS: cimage (an array of chars)
//.....
cimage::cimage()
{
    sized = 0;
    alised = 0;
    allocated_sized = 0;
    allocated_alise = 0;
    dpl = 1;
}

cimage::cimage(int x, int c)
{
    sized = 0;
    alised = 0;
    allocated_sized = 0;
    allocated_alise = 0;
    dpl = 0;
    allocate(x,c);
}

// the problem here's be caused when
// allocate called deallocate
allocated_sized = 0;
allocated_alise = 0;
dpl = 0;
allocate(x,c);
}

// deallocate(cimage)
{
    // Zero out the size
    allocated_sized = 0;
    allocated_alise = 0;
}

// Allocate memory for an cimage from memory
for (int i=0; i<sizeof(cimage); i++)
    pixel[i][3] = img.pixel(i,i);
}

// Deallocate the cimage
{
    deallocate();
}

// Allocate memory for an cimage from memory
void cimage::allocate(int x, int c)
{
    new_C = new_C;
    if ((x > 0) && (c < 0))
        cerr << "Illegal cimage size: " << x << " by " << c << "\n";
    exit(1);
}

// Set cimage sizes
sized = x;
alised = y;
allocated_sized = 0;
allocated_alise = 0;
if ((x != 0) && (c != 0))
}

// Allocate memory for an cimage from outside the class. It will
// Allocate memory if it's larger. If not, it will just adjust the size
void cimage::realloc(int x, int c)
{
    if ((x > allocated_sized) || (c > allocated_alise))
    {
        new_C = allocated_sized;
        new_A = allocated_alise;
        if ((x > new_C) || (c > new_A))
            new_C = x;
            new_A = c;
        allocate(new_C,new_A);
        print("NOTE: IMAGE SIZE REALLOCATED TO (%d,%d)\n", new_C, new_A);
    }
    sized = x;
    alised = c;
}

// Make a cimage with a simple value value - value, x and y are the size
void cimage::fill(char value, int x, int y)
{
}

pixel = (unsigned char **) malloc((unsigned int) sized*sizeof(unsigned char**));
Allocates memory
if (pixel == NULL) // Allocates memory
{
    cerr << "Error: Memory Allocation Error (out of memory)\n";
    exit(1);
}
for (int i = 0; i<sized; i++)
{
    pixel[i] = (unsigned char *) malloc((unsigned int) sized * sizeof(unsigned char));
    Allocates more mem
    if (pixel[i] == NULL)
    {
        cerr << "Error: Memory Allocation Error (out of memory)\n";
        exit(1);
    }
}
allocated_sized = sized;
allocated_alise = sized;
}

// Deallocation memory for an cimage from memory
void cimage::deallocate()
{
    if (allocated_sized != 0) as (allocated_alise != 0)
    {
        for (int i = 0; i < allocated_sized; i++)
            free(char *) pixel[i]);
        if (alised != 0)
            free(char **) pixel);
        sized = 0;
        alised = 0;
        allocated_sized = 0;
        allocated_alise = 0;
    }
}

// Deallocation memory for an cimage from memory
void cimage::deallocate()
{
    if (allocated_sized != 0) as (allocated_alise != 0)
    {
        for (int i = 0; i < allocated_sized; i++)
            free(char *) pixel[i]);
        if (alised != 0)
            free(char **) pixel);
        sized = 0;
        alised = 0;
        allocated_sized = 0;
        allocated_alise = 0;
    }
}

// Allow allocation of memory from outside the class. It will
// Allocate memory if it's larger. If not, it will just adjust the size
void cimage::realloc(int x, int c)
{
    int new_C, new_A;
    if ((x > allocated_sized) || (c > allocated_alise))
    {
        new_C = allocated_sized;
        new_A = allocated_alise;
        if ((x > new_C) || (c > new_A))
            new_C = x;
            new_A = c;
        allocate(new_C,new_A);
        print("NOTE: IMAGE SIZE REALLOCATED TO (%d,%d)\n", new_C, new_A);
    }
    sized = x;
    alised = c;
}

// Make a cimage with a simple value value - value, x and y are the size
void cimage::fill(char value, int x, int y)
{
}

```

【図148】

```

for (i=R1; i<R2; i++) // Copy clipped image
for (j=C1; j<C2; j++)
    temp_pixel[i-R1][j-C1] = pixel[i][j];
allocate(temp_size, temp_sizeC); // Copy temp image to current image
for (i=0; i<sizeC; i++)
    pixel[i][j] = temp_pixel[i][j];
}

// Read in part of a TIFF file
void claimg::readtiff(char *filename)
{
    FILE *tagfile; // Input file
    long int tagfile_position; // Output file position (in bytes)
    long int width; // Width of picture
    long int length; // Length of picture
    long int black; // Value of black pixel
    long int bps; // Bytes per strip
    long int offset; // Start of image data
    long int number; // Tag, type, len, value, // Tag values
    long int i, j;

    if ((tagfile = fopen(filename, "rb")) == NULL)
    {
        cout << "Unable to open file " << filename << "\n";
        exit(1);
    }
    tagfile_position = 0;

    // Read in header and make sure it's Intel
    i = getbytes(tagfile, tagfile_position);
    if (i != 0x4949)
    {
        cout << "... Error: TIFF File not in Intel Format!\n";
        exit(1);
    }

    // Read in Magic number
    i = getbytes(tagfile, tagfile_position);
    if (i != 0x002A)
    {
        cout << "... BAD Magic Number for TIFF file!\n";
        exit(1);
    }

    // Find first file directory
    i = getbytes(tagfile, tagfile_position);
    while(tagfile_position < i)
    {
        i = getbytes(tagfile, tagfile_position);
    }

    // Turn in directory
    number = getbytes(tagfile, tagfile_position);
    while (number > 0)
    {
        tag = getbytes(tagfile, tagfile_position);
        type = getbytes(tagfile, tagfile_position);
        len = getbytes(tagfile, tagfile_position);
        value = getbytes(tagfile, tagfile_position);
        if ((len != 1) && (tag & 32768) == 0)
        {
            cout << "Error: Tag length is 1!\n";
        }
    }
}

```

```

// Make the new claimg (note also adjusted)
for (int i=0; i<sizeC; i++)
    pixel[i][j] = value;
}

// Copy a section of a claimg into another. Image = reference image
// Range is [R1..R2] [C1..C2]
void claimg::clipclaimg(Image, int R1, int R2, int C1, int C2)
{
    int i, j, c;

    if (R2 < R1) // Swap R1 & R2
    {
        i = R1; R1 = R2; R2 = i;
    }
    if (C2 < C1) // Swap C1 & C2
    {
        j = C1; C1 = C2; C2 = j;
    }
    // Check limits
    if (R2 < 0) R2 = 0;
    if (R1 < 0) R1 = 0;
    if (C2 < 0) C2 = 0;
    if (C1 < 0) C1 = 0;
    if (R1 > Image.size() R1 = Image.size();
    if (R2 > Image.size() R2 = Image.size();
    if (C1 > Image.size() C1 = Image.size();
    if (C2 > Image.size() C2 = Image.size();

    allocate(R2-R1, C2-C1);

    for (i=0; i<R2-R1; i++)
    {
        for (j=0; j<C2-C1; j++)
            pixel[i][j] = Image.pixel[i+R1][j+C1];
    }
}

// Clipper: reduce the size of an image by removing the black
// space around it. This adjusts the image size.
void claimg::clipper()
{
    int R1, R2, C1, C2; // The starting and stopping rows of new image
    int i, j;
    struct claimg temp; // Temporary image

    R1 = sizeR; // Initialize to extremes
    R2 = 0;
    C1 = sizeC;
    C2 = 0;

    for (i=0; i<sizeR; i++) // Compute dimensions of new image
    {
        if (pixel[i][0] != 0)
        {
            if (i < R1) R1 = i; // Left edge?
            if (i > R2) R2 = i; // Right edge?
        }
        for (j=0; j<sizeC; j++)
        {
            if (j < C1) C1 = j; // Top edge?
            if (j > C2) C2 = j; // Left edge?
        }
    }

    if (R1 > R2) R1 = R2 = 0; // Check for no pixel found
    if (C1 > C2) C1 = C2 = 0;

    temp.allocate(R2-R1+1, C2-C1+1); // New image is this size
}

```

—339—

```

) // while
{
    while(tagfile_position < raster_offset)
    {
        i = getbytes(tagfile, tagfile_position);
    }
}

//----- READ IN FILE
allocs=((lrc) * length, (int) width);

int v;
for (k = 0; k < length; k++)
{
    for (j = 0; j < width; j++)
    {
        if ((j%8) == 0)
        {
            v = getc(tagfile);
            tagfile_position++;
            if (v == EOF)
            {
                cerr << "\n Error: Premature End of File found in tiff file\n";
                exit(1);
            }
            if ((v & 128) == black*128)
            {
                pixel[j][i] = 0;
            }
            else
            {
                pixel[j][i] = v;
                v = v << 1;
            }
        }
    }
}

fclose(tagfile);

// Read in a byte. File position is updated
// Read in a short. File position is updated
// Read in a long. File position is updated
int change = getbytes(FILE, & long_int file_position, int eoferror)
{
    int o1;
    ci = getc(f);
    if ((ci == EOF) && (eoferror == 1))
    {
        cerr << "\n Error: Premature End of File (in getbytes())\n";
        exit(1);
    }
    file_position++;
    return ci;
}

// Read in an integer from a file. (LW, MSB)
long int change; getbytes(FILE, & long_int fp)
{
    int o1, o2;
    ci = getbytes(f, fp, 1);
    ci = getbytes(f, fp, 1);
    return ci + ci*256;
}

// Read in a long integer
long int change; getbytes(FILE, & long_int fp)
{
}

```

【例 167】

【图 151】

—340—

【図152】

```

// Filename: combine.c
// Routine for combining incorrectly separated characters
//
// =====
#include <charlist.h>
#include <charlist.h>
// Just for debugging!! DELETE EVENTUALLY
#define VERY_SMALL -32000 // A very small integer (for large size)
#define FRAC_VALID_DIST 9.0 // How far apart breaks must be (in % pixels)
// Count the overlapping. Multiply this
// By the smaller dot to get range
#define DOT_TO_JOIN_HEIGHT 1.2 // How much longer the body of the l must be
// than the dot (body > x * dot)
#define PERCENT_PIXEL_JOIN 0.15 // The ratio of the dot the line should
// define PERCENT_PIXEL_JOIN 0.15 // Minimum ratio of dot width / line width
// define PERCENT_PIXEL_JOIN 0.15 // Minimum ratio of dot height / line height
// define PORT_TO_JOIN_SLOPE 0.1 // Ignore top and bottom bit for computing the
// slope of the line in a percent slope.
// define PORT_PERCENT_OVERLAP 0.15 // Multiply this by width of line of s to see
// how far the characters must overlap
// define PORT_DOT_WIDTH_JOIN 5.50 // Minimum ratio of dot width / line width
// define PORT_DOT_HEIGHT_JOIN 1.10 // Minimum ratio of dot height / line height
// define PORT_DOT_JOIN_SLOPE 0.15 // Minimum ratio of dot slope / line slope
// define PORT_JOIN_SLOPE 0.75 // Minimum inverse slope of the line (aka
// inverse of maximum slope of line)
// Define a Combinelist which is a list of 4 characters (with information
// stored about them) that are looked at to see if we need to combine them.
struct Combinelist {
    struct Character *c1; // The first character in the list
    struct Character *c2; // The second character in the list
    // .. (and so on) 10 * some)
    struct Character *c4; // The next character to load in
    int b1; // 1's bit: difference between 1 & 2
    int b2; // 1's bit: difference between 2 & 3
    int d1; // 1's bit: (1 means first higher, 3 means 2nd higher)
    int d2; // The horizontal distance between 1 & 2
    int d3; // (defined from top of 3 to bottom of other)
    int d4; // (VERY_SMALL = undefined)
};

// Define external variables and routines (for displaying and debugging)
extern int cursor_x;
extern int cursor_y;
extern print_character(charline, struct Character *chr, int x, int y);
extern print_charlist(charline, struct Charlist *clist, int x, int y);

// =====
// Character debugging routines
// =====
void debug_print_line(
    struct Charlist *clist, // The charlines
    struct Combinelist *cmb) // The list
{
    int x, y;
    x = y = 10;
}

```

```

// =====
// This routine returns if two characters are above each other.
// Returns 1 for overlapping or character does not exist, 1 for 1st higher,
// 0 for 2nd higher, -1 if also updates h1.
int comb_height_data(
    struct Character *c1, // The first character
    struct Character *c2, // The second character
    int *h1; // The height (updated)
) {
    if (*h1 != -1) // If it's been calculated,
        return (*h1); // don't recalculate it
    if (c1 == 0) // If character doesn't exist
    {
        *h1 = 0; // Return no gap exists
        return 0;
    }
    if (c1->bot <= c2->top) // If bottom of c1 is higher than top of c2
    {
        *h1 = 1; // Return the gap
        return 1;
    }
    if (c2->bot <= c1->top) // If bottom of c2 is higher than top of c1
    {
        *h1 = 2; // Return the gap
        return 2;
    }
    *h1 = 0; // Return no gap exists
    return 0;
}

```

```

for (l=c2-stop - 4; l < c2-stop; l++)
    botpos = c1-curveval2(l) + c2-curveval(l);

// Compute position and exit
// d12 = botpos - toppos;

f = ((float) *d12) / ((float) width_small);
return f;
else
    return 0;

// This routine is called from comdist() with a pair of characters.
// That probably isn't an lsr, j, or i.
// The routine will determine if close enough to be combined.
// The first character
// The second character
// Which character is higher
// The width of the top and bottom
// The width of the smaller character
// Positions of the top and bottoms

int ltoppos, botpos;
int l, i;
float f;

if (h12 == 3)
{
    // Case when c1 is higher than c2
    // Find bottom pos of c1
    if (l12-stop <= 4)
        botpos = c1-xloc;
    else
    {
        botpos = 0;
        botpos += c1-curveval(l) + c1-curveval(l);
    }

    // Find top pos of c2
    if ((c2-stop - c2-stop) < 4)
        toppos = c2-xloc;
    else
    {
        toppos = 0;
        toppos += c2-curveval(l) + c2-curveval(l);
    }

    // Compute position and exit
    // d12 = toppos - botpos;

    f = ((float) *d12) / ((float) width_small);
    return f;
    else
        return 0;
}

// c1 is higher than c1
// Find top pos of c1
if (l12-stop <= 4)
    toppos = c1-xloc;
    else
        toppos = 0;
        toppos += c1-curveval(l) + c1-curveval(l);

// c2 is higher than c2
// Find bottom pos of c2
if ((c2-stop - c2-stop) < 4)
    botpos = c2-xloc;
    else
        botpos = 0;
        botpos += c2-curveval(l) + c2-curveval(l);

// Compute position and exit
// d12 = toppos - botpos;

f = ((float) *d12) / ((float) width_small);
return f;
else
    return 0;
}

// c1 is higher than c1
// Find top pos of c1
if (l12-stop <= 4)
    toppos = c1-xloc;
    else
        toppos = 0;
        toppos += c1-curveval(l) + c1-curveval(l);

// c2 is higher than c2
// Find bottom pos of c2
if ((c2-stop - c2-stop) < 4)
    botpos = c2-xloc;
    else
        botpos = 0;
        botpos += c2-curveval(l) + c2-curveval(l);

// Compute position and exit
// d12 = toppos - botpos;

f = ((float) *d12) / ((float) width_small);
return f;
else
    return 0;
}

```

```

// This routine is called from comb_dists() with a likly 1 or 0
// character. This determines if close enough to be combined.
// This will use the first side of the character to compute closeness
int comb_dists3f
{
    struct char_pos c1,
    struct char_pos c2,
    int h12,
    int c12,
    int width_small,
    int toppos, botpos,
    int t1,
    float f,
    if (h12 == 1)
    {
        // Case when c1 is higher than c2
        // Find bottom pos of c1
        if ((c1->bot - c1->top) < 1)
        {
            // If character very small, use .loc
            botpos = c1->loc;
        }
        else
        {
            botpos = 0;
            for (i=c1->bot - 4; i < c1->bot; i++)
            {
                // botpos == c1->curved[i] + c1->curved[i+1],
            }
        }
        // Find top pos of c2
        if ((c2->bot - c2->top) < 1)
        {
            // If character very small, use .loc
            toppos = c2->loc;
        }
        else
        {
            toppos = 0;
            for (i=c2->top; i < c2->top + 4; i++)
            {
                // toppos == c2->curved[i] + c2->curved[i+1],
            }
        }
        // Compute position and emit
        c12 = toppos - botpos;
        f = ((float) c12) / ((float) width_small);
        return f;
    }
    else
    {
        return 0;
    }
}
// c1 is higher than c2
// Find top pos of c1
// Find top pos of c2
// Find top pos of c12
// toppos = c1->top; // If character very small, use .loc
else
{
    toppos = 0;
    for (i=c1->top; i < c1->top + 4; i++)
    {
        // toppos == c1->curved[i] + c1->curved[i+1],
    }
}
// Find bottom pos of c2
// Find bottom pos of c12
// botpos = c2->bot; // If character very small, use .loc
else
{
    botpos = 0;
    for (i=c2->bot - 4; i < c2->bot; i++)
    {
        // botpos == c2->curved[i] + c2->curved[i+1],
    }
}
}

```


【図154】

```

else
{
    toppos = 0;
    for (i=c1-stop; i < c1-stop + 4; i++)
        toppos += c1-curve2[i] - c1-curve2[i+1];
}
// Find bottom pos of c2
if ((c2-bot - c2-stop) < 4)
    botpos = c2-stop; // If character very small, use .loc
else
{
    botpos = 0;
    for (i=c2-bot - 4; i < c2-bot; i++)
        botpos += c2-curve2[i] - c2-curve2[i+1];
}
// Compute position and exit
*d1 = botpos - toppos;
if (i==(float) *d12) / ((float) w(c2)_small);
if (i < FRAC_VALID_DIST)
    return 1;
else
    return 0;
}

// This routine returns two character heights
// to the caller. It computes characters from top of lower one
// to bottom of higher one. Returns 0 - Bad. 1 - Good. Also updates d12
// Position is generated to be relative to bot height
int comb_dist
{
    struct Character *c1; // the first character
    struct Character *c2; // the second character
    int i; // the first character
    int *d1; // which character is higher
    int *d12; // the last value

    // The size of c1 (width)
    // The size of c2
    // The size of the smaller character
    // The height of c1
    // The height of c2
    float f;

    width_c1 = c1-right - c1-left;
    width_c2 = c2-right - c2-left;
    if (width_c1 < width_c2)
        width_small = width_c1;
    else
        width_small = width_c2;

    if (*d12 != VAR_MAX) // If it's been computed, don't recompute it
    {
        f = ((float) *d12) / ((float) width_small);
        if (f < FRAC_VALID_DIST)
            return 1;
        else
            return 0;
    }

    height_c1 = c1-bot - c1-top;
    height_c2 = c2-bot - c2-top;
    if (h12 > 1) // c1 is higher than c2
}

```

```

// If (height_c1 > DOT_TO_BODY_HEIGHT_1) < height_c2)
return comb_dist_1(c1, c2, h12, d12, width_small);
else
    return comb_dist_2(c1, c2, h12, d12, width_small);
}
// Else c2 is higher than c1
if (height_c2 > DOT_TO_BODY_HEIGHT_2) < height_c1)
{
    return comb_dist_1(c1, c2, h12, d12, width_small);
}
else
    return comb_dist_2(c1, c2, h12, d12, width_small);
}

// This routine sees if larger characters are overlapping enough (left and
// right wise.) used for identifying type 1 percent signs.
// Returns a 1 if overlapping for a type 1 percent, 0 otherwise
// Assumes that the characters ARE overlapping!
int comb_dist_1
{
    struct Character *c1; // The first character
    struct Character *c2; // The second character

    int i1, i2; // The length of the characters
    int a1, a2; // The left sides of the character
    int b1, b2; // The right sides of the character
    int swap; // Misc var

    i1 = c1-right - c1-left;
    i2 = c2-right - c2-left;
    if (i1 > i2)
    {
        swap = c1;
        c1 = c2;
        c2 = swap;
        i1 = i2;
        i2 = swap;
    }
    // c1 is longer
    // Swap to make c1 shorter

    if ((float) i1 / (float) i2 < PERCENT_TYPE1_MIN)
        return 0;
    if ((float) i1 / (float) i2 > PERCENT_TYPE1_MAX)
        return 0;

    if (c2-left < c1-left)
        a1 = c2-left;
        a2 = c1-left;
    if (c2-right > c1-right)
        a2 = c1-right;
        a1 = c2-right;
    if ((float) (a2 - a1) / ((float) i1) > PERCENT_TYPE1_MAX)
        return 1;
    else
        return 0;
}

```


【図156】

```

if (c1->left < c2->left)
    left = c1->left;
else
    left = c2->left;
if (c1->right > c2->right)
    right = c1->right;
else
    right = c2->right;

// Load information about the stuff.
c1->type = c1;
c1->top = top;
c1->bot = bot;
c1->left = left;
c1->right = right;
c1->next = c2->next;
if (c1->next == 0)
    c1->next = c2;
c1->type = CHARACTER_COINED;
c2->type = CHARACTER_ERASED;
}

// This routine will combine characters 1 and 2, and slide them
// out
void comb_combine12()
{
    struct comblist *comb; // The list
    comb_combine(comb->c1, comb->c2);
    comb_slide(comb);
}

// The main routine
// This is a recursive function to combine a charlist
void charlist_combine()
{
    struct charlist *clist; // Image of the line
    clargest_line;
    struct comblist *comb; // The 4 characters in the que we're evaluating
    int dummy;

    // Initialise the combination list
    comb->next = clist->first;
    comb_slide(comb);
    comb_slide(comb);
    comb_slide(comb);
    comb_slide(comb);

    // Main loop: Warning: indentation is not standard (but not too bad)
    while(comb->c2 != 0) // Loop until list is empty
    {
        // Combine 1 & 2
        if (comb->height_delta(comb->c1, comb->c2, &comb->h12))
            if (comb->dist(comb->c1, comb->c2, &comb->d12))
            {
                // Combine 3 & 3? Instead!
                if (comb->c3 != 0)
                {
                    if (comb->height_delta(comb->c2, comb->c3, &comb->h23))
                        if (comb->dist(comb->c2, comb->c3, &comb->d23))
                        {

```


[illegible]

[illegible]

【図160】

```

//def PRINT_INFO
int dummy;
endif

histogram = vector<line>::size(); // Allocate the memory
while (charen != 0) // Loop until end of list
{
    cd_char(line, clist, lcharon, histogram);
} //def PRINT_INFO
printf("Enter a to continue\n");
scanf("%d", &dummy);
}
charen = charen->next; // Get the next character
}
free_vector(histogram); // Deallocate the memory
}

// Check for the dash between a character, and break it if needed
void cd_char(
// The line
// Point to charlist
// Point to the character we're on (modified
// to skip past newly inserted divisions)
// The histogram (previously allocated)
int *hist)
{
    int p;
    // Offset if p[0] width is too small.
    if ((ch->right - (ch->left) < MIN_ABS_WIDTH_TO_CONSIDER)
        return;
} //def PRINT_INFO
re_char();
printf("%c\n", line, *ch, 10);
re_char();
printf("Character info:\n");
printf(" top/bot = (%d, %d)\n", (ch->stop, (ch->bot));
printf(" left/right = (%d, %d)\n", (ch->left, (ch->right));
endif

// Compute the histogram
int histogram(line, ch, hist)
{
    // Break left side
    // cd_char dash out(line, ch, hist);
    if (p != 0)
    {
        // Cut the character
        // Get the histogram
        // Get the histogram
        printf("left dash out\n");
        printf("Character info:\n");
        printf(" top/bot = (%d, %d)\n", (ch->stop, (ch->bot));
        printf(" left/right = (%d, %d)\n", (ch->left, (ch->right));
        endif
    }
    // Break right side
    p = cd_char_dash_out(line, ch, hist);
    if (p != 0)
    {
        // Cut the character
        //def PRINT_INFO
        printf("right dash out\n");
        endif
    }
}

// MAIN ROUTINE
// This will cut the dashes of characters
// The change
// The change
// The current charlist
// The area of the histogram
// The character we're processing

```

【図161】

```

loc = loc_count + 0; // for position of the char (on the fly)
for (t=top; t<bot; t++)
{
    while (curvel[t] < curvel[t+1]) // Loop until they're equal
    {
        if ((line.pixel[t](curvel[t]) & 1)) // If there is a non-empty space
            // to the right of the left border
            break;
        else
            // Then exit left border not moved
            break;
    }
    while (curvel[t] > curvel[t+1]) // Do the same with the right border
    {
        if ((line.pixel[t](curvel[t+1]) & 1))
            break;
        else
            curvel[t] += 1;
    }
    if (curvel[t] != curvel[t+1]) // there is a pixel on this line...
    {
        loc += curvel[t] + curvel[t+1];
        loc_count++;
    }
    if (loc_count != 0)
    {
        loc = loc + 4 * loc_count; // compute the position (used for loc)
        pos = loc/8; // Compute the real position
    }
    // Find left and right
    left = curvel[top];
    right = curvel[bot];
    for (t=top; t<bot; t++)
    {
        if (curvel[t] < left)
            left = curvel[t];
        if (curvel[t] > right)
            right = curvel[t];
    }
    if (loc == 0)
    {
        loc = 4 * (left+right);
        pos = (left+right)/2;
    }
    for (t=top; t<bot; t++)
    {
        if (curvel[t] != curvel[t+1]) // Fill in the blank lines with pos
        {
            curvel[t] = curvel[t+1] + pos;
        }
        // copy all the new attributes in
        (*ch)->loc = loc;
        (*ch)->top = top;
        (*ch)->bot = bot;
        (*ch)->left = left;
        (*ch)->right = right;
    }
}

// This routine will cut "ch" into two at point p. "clat" is updated.
// void l32_cut_character(
//     struct charlist *clat, // The character list
//     struct charlist *ch, // The character to be cut
//     int p) // Where to cut the character
// {

```

```

// filename: level3.c
// This file contains the routines to do level 3 segmentation (slightly
// modified) using method 3 (histogram approach)
//
// The includes
//
// #include "charlist.h"
// #include "changes.h"
// #include "math.h"
//
// The constants
//
// #define MIN_CHAR_THRESHOLD 10
//
// This routine will adjust left, right, curvel, top, bottom
// and pos to represent the character. It will then
// find the left and right borders of the character.
// Possible optimization: usually, this is called when a character is
// split, so shrinking one side is 'in val'.
// void l32_shrink_char(
//     struct charlist *ch) // The line
// {
//     int top, bot, left, right; // Adjust to the character we're on
//     int t; // The new left and right stuff.
//     int loc, loc_count, pos; // pos = position, loc = pos * 8
//     int left, right, curvel; // Short-hand for "ch->curvel" and 2
//     int flag;
//
//     curvel = (*ch)->curvel;
//     curvel = (*ch)->curvel;
//
//     // Adjust the top of the character. If it needs to move down
//     flag = 0;
//     for (t=top; t<bot; t++)
//     {
//         if (curvel[t] < curvel[t+1])
//             if (line.pixel[top](t) & 1) // Stop at the first on pixel
//             {
//                 flag = 1;
//                 break;
//             }
//     }
//
//     // Adjust the bottom of the character
//     for (t=bot; t>top; t--)
//     {
//         if (curvel[t] > curvel[t+1])
//             if (line.pixel[bot](t) & 1) // Stop at the first on pixel
//             {
//                 flag = 1;
//                 break;
//             }
//     }
//
//     // Now adjust curvel and curvel. Updating left and right to 'shrink'.
//     // tightly around the border of the character (move the borders up to the
//     // edge of the character.

```


【図162】

```

int min, min_at;
int i;
int o;

// Where the minimum value is
// Left and right boundaries of where to search

// Compute the histogram
132_histogram[line, ch, hist];

p = ('ch' - 'left');
while (p < p_end)
{
    // start at left side
    // Exit when we've scanned the entire histogram
    while (p < p_end)
    {
        // Looking for top-bottom crossing of threshold
        if (hist[p] < historesis_threshold)
        {
            break;
        }
        // This is the left side of the area to search
        1 = p;
        while (p < p_end)
        {
            // Looking for bottom-top crossing of threshold
            if (hist[p] > historesis_threshold)
            {
                break;
            }
            // Remember right edge;
            r = p;
            if (p < p_end)
            {
                // If we haven't gone over the edge
                min = hist[1];
                min_at = 1;
                for (i = 1; i < r; i++)
                {
                    if (hist[i] < min)
                    {
                        min = hist[i];
                        min_at = i;
                    }
                }
            }
            // If the min is small enough, cut there
            {
                min = min_cut_threshold;
                132_cut_character[line, clist, ch, min_at]; // Cut the character
                // Note: ch is changed to the right char,
                // which shouldn't mess things up
                printf("Level 3 made a cut: %d\n", i);
            }
        }
    }
}

// *****
// MAIN ROUTINE
// *****
// This will do histogram touching on the characters for level 2 method 3
void level2(cimgapi clist,
            struct Charlist* clist) // The image
{
    int histogram; // The area of the histogram
    struct Character* charon; // The character we're processing
    histogram = (vector<line>::size()); // Allocate the memory
    charon = clist->first;
    while (charon != 0)
    {
        132_insert_break(line, clist, charon, histogram, MIN_CUT_THRESHOLD);
    }
}

```

```

struct Character* newchar; // The new character
int i;

newchar = (struct Character*) malloc(sizeof(struct Character));
// of the charlist (the next blank character)

for (i = 0; i < clist->vector::size(); i++) // Loop through the char (up-down)
{
    newchar->curve2[i] = ('ch' - 'curve2[i]');
    newchar->curve2[i] = ('ch' - 'curve2[i]'); // Break the character
    if (newchar->curve2[i] < p) // Make sure curves of new char are > p
    {
        newchar->curve2[i] = p;
    }
    if (newchar->curve2[i] > p) // Make sure curves of old char are <= p
    {
        ('ch' - 'curve2[i]') = p;
    }
    if (newchar->curve2[i] > p)
    {
        newchar->top = ('ch' - 'top');
        newchar->bot = ('ch' - 'bot');
        newchar->left = ('ch' - 'left');
        newchar->right = ('ch' - 'right');
        ('ch' - 'right') = p;
    }
    132_subincomp_char(line, newchar); // Place newchar into the charlist
    Charlist_place_new_char(clist);
    132_subincomp_char(line, ch);
    Charlist_replace_char(clist, 'ch');
    *ch = newchar; // Adjust ch to the new character
}

// *****
// Compute the histogram of a character
// *****
void 132_histogram(
    struct Character* ch,
    struct Character* ch,
    int* hist)
{
    int i;
    // Save the histogram
    for (i = 'ch' - 'left'; i < ('ch' - 'right'); i++)
    {
        hist[i] = 0;
    }
    for (i = 'ch' - 'top'; i < ('ch' - 'bot'); i++) // Compute histogram
    {
        for (j = ('ch' - 'curve2[i]'); j < ('ch' - 'curve2[i] + 1); j++)
        {
            hist[i]++;
        }
    }
}

// *****
// Check if a character needs to be broken, and break it if so
// *****
void 132_insert_break(
    struct Character* clist,
    struct Character* ch,
    int* hist,
    int min_cut_threshold,
    int historesis_threshold)
{
    int p;
    int p_end;
    // The line
    // Pointer to the character we're on (modified)
    // to skip past newly inserted divisions)
    // The histogram (previously allocated)
    // The minimum value to make a cut
    // historesis_threshold (min_cut_threshold)
    // The cursor
    // The end of the cursor
}

```

—352—

[illegible]

1

—353—

[illegible]

—355—

[illegible]

[illegible]

—358—

```

allocated_s, & max_s, index_0s);

// Find maximum projection
max_value_s = a(0);
for (j=1; j<max_c; j++)
    if (all) > max_value_s
        max_value_s = a(j);

// Choose range to ignore points before first peaks on left and right
z1 = 0;
z2 = max_c-1;
if (tcheta != 0)
{
    // Find left peak while slope is still rising), and above LA_IGNORE_HEIGHT
    while ((a(z1) < a(z1+1)) || (a(z1) < a(z1+2))) { z1++; }
    // If got to left edge
    if ((z1-z2) == LA_IGNORE_HEIGHT*max_value_s)
    {
        min_s = min_at = -1;
        free_vector(s);
        return;
    }

    // Return no minimum found
}

// Find right peak while slope is falling (from right to left)
while ((a(z2) < a(z2-1)) || (a(z2) < a(z2-2))) { z2--; }
a(z1) = la_ignore_height*max_value_s;
a(z2) = la_ignore_height*max_value_s;

if ((z2 <= z1) || (z2 <= 0))
{
    min_s = min_at = -1;
    free_vector(s);
    return;
}

find_minimum_in_vector(s, z1, z2, min_s, min_at);

// ***** The above is mostly copied from void find_min_valley_at_cheta ....
// ***** The following section is VERY VERY poorly written!!!!!!

for (icursor_col_left=1; icursor_col_right=1; i++) // Draw line above cursor2_y at s
{
    // Plot the image
    print_image_with_breaks(image, bist, bnum1, bnum2, cursor2_y);

    // Find position of the C line and plot it
    struct Break bi;
    break_list_search(breaklist, bnum1, abbi); // Find initial offset
    int cx = cursor2_y - bi.col/image_sizef + c * 2; // Get first break
    int cy = cursor2_y - bi.col/image_sizef + c * 2;
    plot_line_at_angle(Cx, Cy, cx, cursor2_y - 4, Cy + S); // Right line

    cursor2_Y -- image_sizef + 544;

    // Plot the histogram
    int cntemp = simple_histogram(hist, max_c, cursor2_X, cursor2_Y);
    printf("Histogram: %d,%d", histy, cntemp);
    if (tcheta != 0)
    {
        int ztt = (int) (LA_IGNORE_HEIGHT * (double) max_value_s);

```


【圖 172】

```

// Find p2, the valley to the right of the given peak
p1x = find_valley_to_right(p1x, proj_size, p1x, max_valley_depth);
if (p1x == -1) // If not found, make it the right side
    p1x = proj_size - 1;
p1y = proj(p1x);

// Find p4, the next peak to the right of valley p2
p1x = find_peak_to_right(p1x, proj_size, p1x, min_peak_height);
if (p1x == -1) // If not found, make it the right side
    p1x = proj_size - 1;
p1y = proj(p1x);

// Find p3, the nearest valley to the left of peak at p4
p1x = find_valley_to_left(p1x, p1x, max_valley_depth);
if (p1x == -1) // If not found, make it the left side
    p1x = 0;
p1y = proj(p1x);

thet1 = atan(double) (p1x-p1x) / (double) (p1y-p1y); // This is positive
thet2 = atan(double) (p1x-p1x) / (double) (p1y-p1y); // This is negative
}

// ===== Makhadm's optimal threshold finding routine =====
void compute_optimal_threshold()
{
    int proj; // The VPP (in atan)
    int p1x; // The size of proj
    int min_valley_depth; // The size of proj
    int min_peak_height; // The threshold level
    int max_valley_depth; // The output
    int i;
    double a, b;
    FILE *fp;

    for (i = 0; i < proj; i++)
    {
        int max_valley_depth;
        int i;
        double a, b;
        FILE *fp;

        if ((fp = fopen("constants.dat", "r")) == NULL)
            cerr << "Error in opening constants.dat\n";
            exit(1);
    }

    fscanf(fp, "%lf", &a);
    for (i = 0; i < proj; i++)
    {
        fscanf(fp, "%lf", &a, &b);
        if ((a < -1) || (b < -1))
            // Read a line
            // Exit
            // Return an error
            min_valley_depth = -1;
            min_peak_height = -1;
            return;
    }
    fclose(fp); // Close the file

    // printf("... LEVEL ed = %f, atan", level, a, b);
    // Find maximum projection (for find peak and valley heights:
    if (proj_size == 0)
    {
        max_valley_depth = 0;
        min_valley_depth = 0;
        return;
    }
}

// Set to not found
peak = -1;
while ((i-pos == 0) && (peak == -1))
    if ((proj[pos]) <= proj[pos+1]) i = proj[pos+1] == min_peak_height;
    peak = pos + 1;
return peak;
}

// Subroutine to find next valley to the right in a projection
// -1 = none found
int find_valley_to_right()
{
    int proj_size; // The projection
    int p1x; // The size of the projection (not inclusive)
    int p1y; // Where to start looking
    int max_valley_depth; // Highest point a valley can be
    int valley; // Where the valley is

    while (i-pos <= 0) // Set to not found
        while ((proj[pos+1] <= proj[pos]) && (valley == -1))
            if ((proj[pos+1] <= proj[pos]) && (proj[pos+1] <= max_valley_depth))
                valley = pos + 1;
    return valley;
}

// Subroutine to find next valley to the left in a projection
// -1 = none found
int find_valley_to_left()
{
    int proj_size; // The projection
    int p1x; // Where to start looking
    int max_valley_depth; // Maximal value for the valley
    int valley; // Where the peak is

    while (i-pos >= 0) // Set to not found
        while ((i-pos == 0) && (valley == -1))
            if ((proj[pos] <= proj[pos+1]) && (proj[pos+1] <= max_valley_depth))
                valley = pos + 1;
    return valley;
}

// ===== Compute the Makhadm's angles given a valley position =====
void find_valley_angles()
{
    int proj; // The projection vector
    int proj_size; // The size of the projection vector
    int valley_pos; // The position of the right of the 1st peak
    int valley_pos; // The position of the left of the 1st peak
    double thet1; // Angle of right wall
    double thet2; // Angle of left wall
    int min_peak_height; //
    int max_valley_depth;

    // The coord of left peak
    // The coord of left valley bottom
    // The coord of right valley
    // The coord of right peak
    int p1x, p1y;
    int p1x, p1y;
    int p1x, p1y;
    int p1x, p1y;

    // Find p1, the peak to the left of the valley
    p1x = find_peak_to_left(proj, valley_pos, min_peak_height);
    if (p1x == -1) // If not found, make it the left side
        p1x = 0;
    p1y = proj(p1x);
}

```

—360—

[illegible]

[illegible]

【图 1 7 5】

```

// cout << " At theta3 = 0, min = " << min << "\n";
insert_angular_break(image, blist, start_breakbreak_count,
stop_breakbreak_count, min_theta, min_beta, min_alpha);

// Check theta1 = 0
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta1-min_degrees, tmin, tmax);
cout << " At theta1 = 0, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta1-min_degrees);

// Check theta1 = 90
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta1+min_degrees, tmin, tmax);
cout << " At theta1 = 90, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta1+min_degrees);

// Check theta1 = 45
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta1-min_degrees, tmin, tmax);
cout << " At theta1 = 45, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta1-min_degrees);

// Check theta1 = 135
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta1+min_degrees, tmin, tmax);
cout << " At theta1 = 135, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta1+min_degrees);

// Check theta2 = 0
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta2-min_degrees, tmin, tmax);
cout << " At theta2 = 0, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta2-min_degrees);

// Check theta2 = 90
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta2+min_degrees, tmin, tmax);
cout << " At theta2 = 90, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta2+min_degrees);

// Check theta2 = 45
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta2-min_degrees, tmin, tmax);
cout << " At theta2 = 45, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta2-min_degrees);

// Check theta2 = 135
find_min_value_at_theta(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, theta2+min_degrees, tmin, tmax);
cout << " At theta2 = 135, min = " << min << "\n";
insert_angular_computations(min_alpha, min_beta, min_theta, min_alpha,
theta2+min_degrees);

// Insert the best break
// cout << " Final at " << min_alpha << ", theta = " << min_theta << ", beta = " << min_beta << "\n";
min_alpha = min_alpha;
min_beta = min_beta;
min_theta = min_theta;
insert_angular_break(image, blist, start_breakbreak_count,
stop_breakbreak_count, C, min_alpha, min_beta, min_theta);
}

// cout << " At theta3 = 90, min = " << min << "\n";
insert_angular_break(image, blist, start_breakbreak_count,
stop_breakbreak_count, min_beta, min_theta, min_alpha);

break_count++;
} // Move i to the next peak do another pass
} // }

////////////////////////////////////
// This routine is used as a hint to pass a character to recognition
// so we can go ahead or stop depending on what it returns. It returns a
// 1 to go ahead, -1 otherwise (if zero, it's backward).
////////////////////////////////////
int pass_to_recognition_level() {
    // The image
    image = image;
    // The breaklist
    breaklist = blist;
    int bi;
    int b1;
    int b2;
    int *a;
    // The matrix to send to the recognition routine
    int size_c; // The size of matrix c
    int i;
    int j;
    int k;
    int array1;
    int array2;
    array1 = ivector(image.size);
    array2 = ivector(image.size);
    breaklist_return_values(blist, b1, b2, array1, array2);

    z = -1;
    for (i=0; i<(image.size); i++)
        if (array1[i] < 1) // Out range
            i = array1[i];
        if (array2[i] < 1)
            i = array2[i];
        if (array1[i] > 1)
            i = array1[i];
        if (array2[i] > 1)
            i = array2[i];
        size_c = z + 1;
        image = image.size;
        size_c = image.size;
        if (size_c > MAX_CHARACTERS) // If too wide, automatic failure
            return 0;
        a = matrix(size_c, size_c);
        for (i=0; i<(size_c); i++)
            for (j=0; j<(size_c); j++)
                a[i][j] = 0;
        for (i=0; i<(image.size); i++)
            for (j=array1[i]; j<array2[i]; j++)
                a[i][j-1] = 1;
        return_value = recognition(a, size_c, size_c);
        free(matrix);
        return return_value;
}

```


[1 7 8]

```

    printf("Combining (N M) and passing deeper\n");
    breaklist_remove_breaklist(bn1);
    breaklist_remove_breaklist(bn2);
    bn2 = breaklist_return_break_ptr(blist, bn1);
    bck = breaklist_return_break_ptr(blist, bn1);
    bck->next_recog = -1;
    b == layer_goodout(image, blist, bn, bn+1, thresh_level, 1);
    return b;
}
if ((R(bn+3) == 0) && (R(bn+3) == 1)) // YNNY
{
    printf("Y N M Y N\n");
    printf("Trying Y (N M) - ");
    if (pass_to_deeper_level(image, blist, bn+1, bn+4))
    {
        printf("Good\n");
        breaklist_remove_breaklist(blist, bn+2);
        breaklist_remove_breaklist(blist, bn+3);
        bn+2; breaklist_return_break_ptr(blist, bn+1);
        bck = breaklist_return_break_ptr(blist, bn+1);
        bck->next_recog = 1;
        return b;
    }
    printf("Combining (N M Y) and passing deeper\n");
    breaklist_remove_breaklist(blist, bn+1);
    breaklist_remove_breaklist(blist, bn+2);
    bn+2; breaklist_return_break_ptr(blist, bn1);
    bck = breaklist_return_break_ptr(blist, bn1);
    bck->next_recog = -1;
    b == layer_goodout(image, blist, bn, bn+1, thresh_level, 1);
    return b;
}
}
printf("WARNING: No pattern matched the goodcut routines\n");
return b;
}

// Master Segmentation Routine
// Valid input image
// The input image, a gray image
// Space allocated for a break list
// This will be initialized
breaklist = blist;
int first_break, last_break, // Misc. Used for break numbers
int b; proj_alloc, // Projection vector and its size
struct Break *brk;

// Allocate the projection vector
proj_alloc = image_alloc;
proj = vector(proj_alloc);

// Find first and last breaks
first_break = 0;
last_break = blist->blist_count-1;
if (last_break != 1)

```

```

    printf("Combining (N M) and passing deeper\n");
    breaklist_remove_breaklist(bn1);
    breaklist_remove_breaklist(bn2);
    bn2 = breaklist_return_break_ptr(blist, bn1);
    bck = breaklist_return_break_ptr(blist, bn1);
    bck->next_recog = -1;
    b == layer_goodout(image, blist, bn, bn+1, thresh_level, 1);
    return b;
}
if ((R(bn+3) == 0) // PPNY
{
    printf("Y N M Y N\n");
    printf("Trying Y (N M) - ");
    if (pass_to_deeper_level(image, blist, bn+2, bn+4))
    {
        printf("Good\n");
        breaklist_remove_breaklist(blist, bn+3);
        bn+3; breaklist_return_break_ptr(blist, bn+2);
        bck = breaklist_return_break_ptr(blist, bn+2);
        bck->next_recog = 1;
        return b;
    }
    printf("Bad\n Trying Y (N M) - ");
    if (pass_to_deeper_level(image, blist, bn, bn+4))
    {
        printf("Good\n");
        breaklist_remove_breaklist(blist, bn+2);
        breaklist_remove_breaklist(blist, bn+3);
        bn+3; breaklist_return_break_ptr(blist, bn+1);
        bck = breaklist_return_break_ptr(blist, bn+1);
        bck->next_recog = -1;
        return b;
    }
    printf("Bad\n Trying Y (N M) - ");
    if (pass_to_deeper_level(image, blist, bn, bn+4))
    {
        printf("Good\n");
        breaklist_remove_breaklist(blist, bn+2);
        breaklist_remove_breaklist(blist, bn+3);
        bn+3; breaklist_return_break_ptr(blist, bn+1);
        bck = breaklist_return_break_ptr(blist, bn+1);
        bck->next_recog = -1;
        return b;
    }
    printf("Bad\n Combining Y (N M) and passing to deeper level\n");
    breaklist_remove_breaklist(blist, bn+2);
    breaklist_remove_breaklist(blist, bn+3);
    bn+3; breaklist_return_break_ptr(blist, bn+1);
    bck = breaklist_return_break_ptr(blist, bn+1);
    bck->next_recog = -1;
    b == layer_goodout(image, blist, bn+1, bn+2, thresh_level, 1);
    return b;
}
if ((R(bn+3) == 1) // We have YNY
{
    printf("Y N M Y N\n");
    printf("Trying Y (N Y) - ");
    if (pass_to_deeper_level(image, blist, bn+1, bn+3))
    {
        printf("Good\n");
        breaklist_remove_breaklist(blist, bn+2);
        bn+2; breaklist_return_break_ptr(blist, bn+1);
        bck = breaklist_return_break_ptr(blist, bn+1);
        bck->next_recog = 1;
        return b;
    }
}

```


—367—

—368—

[illegible]

【図182】

```

1: if (fp = fopen("constants.dat", "r")) == NULL)
2:   cerr << "underror in opening constants.dat\n";
3:   exit(1);
4:
5: // *** This should change before the final version. When we know the size
6: // of constants.dat, hardcode it!
7: a = vector(200);
8: b = vector(200);
9:
10: for (i=0; i<200; i++)
11: {
12:   fscanf(fp, "%f", &a[i], &b[i]);
13:   if ((a[i] < 0) || (b[i] < 0))
14:   {
15:     max_thresh_level = i;
16:     break;
17:   }
18: }
19:
20: if (i==200)
21: {
22:   print("**** Error: only 200 entries in constants.dat is allowed! Change the size
23:   'n'");
24:   exit(1);
25: }
26:
27: Const1 = fvector(max_thresh_level);
28: Const2 = fvector(max_thresh_level);
29: for (i=0; i<max_thresh_level; i++)
30: {
31:   Const1[i] = a[i];
32:   Const2[i] = b[i];
33: }
34:
35: free_fvector(a);
36: free_fvector(b);
37:
38: // =====
39: // (this must be called before calling compute_rotated_projection)
40: //
41: void k4_initialize_rot()
42: {
43:   int d_index;
44:   double d_t_d;
45:   int N;
46:   int a,b;
47:
48:   d = ROT_RANGE;
49:   N = MAX_CHAR_LENGTH;
50:   Rot = matrix(a, b);
51:
52:   for (d = ROT_MIN_RESOLUTION; d<ROT_MAX_RANGE; d+=ROT_MIN_RESOLUTION)
53:   {
54:     d_index = (int) ((d-ROT_MIN_RESOLUTION+.5) / 1);
55:     for (i=0; i<N; i++)
56:     {
57:       for (j=0; j<N; j++)
58:       {
59:         Rot[d_index][i] = (int) ((double) z1-.5);
60:       }
61:     }
62:   }
63: }
64:
65: // ===== PLOTTING ROUTINES =====
66: void k4_plot_char(
67:   struct Image
68:   struct Character *ch, // The character to plot
69:   int x, // The upper left hand corner i.left, .top; to plot
70:   int y)
71: {
72:   int i,j;
73:   for (i=ch->top; i < ch->bot; i++)
74:   {
75:     for (j = ch->curved[i]; j < ch->curved[i+1]; j++)
76:     {
77:       if (image.pixel[i][j] < 1)
78:       {
79:         x = plot_x(image->ch->curved[i], j-ch->left, y+i);
80:         x = plot_x(image->curved[i+1]-ch->left, y+i);
81:         x_flush();
82:       }
83:     }
84:   }
85:
86:   void k4_plot_charlist(
87:     char *line, // The image
88:     struct Charlist *clist, // The charlist
89:     int x, // The position
90:     int y)
91:   {
92:     struct Character *ch;
93:     int i, start;
94:     for (i = 0; i < clist->n; i++)
95:     {
96:       ch = clist->list[i];
97:       while (ch != 0)
98:       {
99:         if (x + ch->right - ch->left > WINDOW_SIZE_X - 20) // Wrap?
100:         {
101:           x = x_start;
102:           y = y_start + 15;
103:         }
104:         k4_plot_char(line, ch, x, y);
105:         x = ch->right - ch->left;
106:         ch = ch->next; // Get next character
107:       }
108:     }
109:   }
110:
111: // This routine will plot a character in the lower left hand corner
112: // (for the purpose of displaying to prompt for recognition)
113: void k4_plot_lower_left(
114:   struct Image, // The image
115:   struct Character *ch)
116: {
117:   int i,j;
118:   int x_start;
119:   int y_start;
120:   for (i=0; i<ch->bot - ch->top; i++)
121:   {
122:     for (j=0; j<ch->right - ch->left; j++)
123:     {
124:       x = WINDOW_SIZE_X - 15;
125:       y = WINDOW_SIZE_Y - 15;
126:       x_flush();
127:     }
128:   }
129:   for (i=ch->top; i < ch->bot; i++)
130:   {
131:     for (j=ch->curved[i]; j < ch->curved[i+1]; j++)
132:     {
133:       if (image.pixel[i][j] < 1)
134:       {
135:         x = plot_x(image->ch->curved[i], j-ch->left, y+i);
136:         x = plot_x(image->curved[i+1]-ch->left, y+i);
137:         x_flush();
138:       }
139:     }
140:   }
141:   for (i=0; i<ch->right - ch->left; i++)
142:   {
143:     for (j=0; j<ch->bot - ch->top; j++)
144:     {
145:       x = WINDOW_SIZE_X - 15;
146:       y = WINDOW_SIZE_Y - 15;
147:       x_flush();
148:     }
149:   }
150: }

```

【図 183】

[illegible]

【図184】

```

int *ang_proj_min, // The min index of ang_proj[] (inclusive) (output)
int *ang_proj_max, // The max index of ang_proj[] (exclusive) (output)
// Note: projection put in proj[]

int r,c; // Looping variables

// Do a standard histogram
if (angle == 0) // Do a standard vpp histogram
{
    *ang_proj_min = range[];
    if (*ang_proj_min < ch->left)
        *ang_proj_min = ch->left;
    *ang_proj_max = ch->right;
    if (*ang_proj_max > ch->right)
        *ang_proj_max = ch->right;

    for (c=ch->left; c<ch->right; c++) // Zero projection array
        ang_proj[c] = 0;
    for (c=ch->left; c<ch->right; c++) // Compute the projection
        for (i=ch->curveval[i]; i<ch->curveval[i+1]; i++)
            if (image.pixel[i] & 1)
                ang_proj[c]++;
    return;
}

// Do a theta is greater than zero
if (angle > 0) {
    // Get the index to use in rot
    *ang_proj_min = range[];
    if (*ang_proj_min < ch->left)
        *ang_proj_min = ch->left;
    *ang_proj_max = range[];
    if (*ang_proj_max > ch->right)
        *ang_proj_max = ch->right;
    *ang_proj_min = ch->left;
    *ang_proj_max = rot(angle)[ch->bot-ch->top];
    for (c=ang_proj_min; c<ang_proj_max; c++) // Zero projection array
        ang_proj[c] = 0;
    for (c=ch->left; c<ch->right; c++) // Compute the projection
        for (i=ch->curveval[i]; i<ch->curveval[i+1]; i++)
            if (image.pixel[i] & 1)
                (ang_proj[(c-rot(angle)[ch->bot-r-1])])++;
    return;
}

// Do a theta is less than zero
if (angle < 0) {
    // Get the index to use in rot
    *ang_proj_min = range[];
    if (*ang_proj_min < ch->left)
        *ang_proj_min = ch->left;
    *ang_proj_max = range[];
    if (*ang_proj_max > ch->right)
        *ang_proj_max = ch->right;
    *ang_proj_min = range[];
    if (*ang_proj_max < ch->right)
        *ang_proj_max = ch->right;
    *ang_proj_max = rot(angle)[ch->bot-ch->top];
    for (c=ang_proj_min; c<ang_proj_max; c++) // Zero projection array
        ang_proj[c] = 0;
    for (c=ch->left; c<ch->right; c++) // Compute the projection
        for (i=ch->curveval[i]; i<ch->curveval[i+1]; i++)
            if (image.pixel[i] & 1)
                (ang_proj[(c-rot(angle)[ch->bot-r-1])])++;
    return;
}
}

// eliminate white space from the left and right side of a character
// after it is out so the size won't be too inaccurate.
// Call when pixels the out, ch->left char, ch->right char
void trim_image(char *ch)
{
    struct Character *ch; // The left character (only does right side)
    struct Character *ch; // The right character (only does left side)

    int r,c; // The new edges of the characters
    int new_left, new_right;
    // right side of char! Look for 1st on pixel
    for (c = ch->right-1; c>ch->left; c--)
        for (r = ch->top; r<ch->bot; r++)
            if (ch->curveval[r] & 1) // if pixel is on
                if (image.pixel[r][c] & 1) // if pixel is on
                    new_right = c;
    // Break out of the loop
    goto mit_loop;
}

// set right side of character to this value
ch->right = new_right;
for (r=ch->top; r<ch->bot; r++)
    if (ch->curveval[r] > new_right)
        ch->curveval[r] = new_right;
    if (ch->curveval[r] > new_right)
        ch->curveval[r] = new_right;
}

// Do the left side of ch
// Look for the 1st on pixel
for (c=ch->left; c<ch->right; c++)
    for (r=ch->top; r<ch->bot; r++)
        if ((c==ch->curveval[r]) & 1) // if ch->curveval[r] & 1)
            if (image.pixel[r][c] & 1) // if pixel is on
                new_left = c;
    // Break out of the loop
    goto mit_loop;
}

// set the left side of character to this value
ch->left = new_left;
for (r=ch->top; r<ch->bot; r++)
    if (ch->curveval[r] < new_left)
        ch->curveval[r] = new_left;
    if (ch->curveval[r] < new_left)
        ch->curveval[r] = new_left;
}

//..... THE ROTATED PROJECTION ROUTINES .....
void rot_projection(char *ch) // The input image
{
    struct Character *ch; // The character to find the projection on
    int angle; // The rotation/ROT_ANGLE_RESOLUTION must be within a
    // valid range!
    // The column range to search through
    int range;
    int range;
    int *ang_proj; // The matrix to put the projection into (*MUST*
    // be previously allocated)

```

【☒ 1 8 5】

—372—

【図186】

```

// The projection
int proj_site,
// The y-axis of the projection (not inclusive)
// Where to start looking
int max_valley_depth // Highest point a valley can be
{
    int valley // Where the valley is.
    // Set to not found
    while ((valley < proj_min) || (valley >= proj_max))
    {
        if ((proj[pos-1] <= proj[pos]) || (proj[pos+1] <= max_valley_depth))
            valley = pos - 1;
        return valley;
    }
}

// Subroutine to find next valley to the left in a projection
// -1 = none found
int find_valley_left(
    int proj_min, // The leftmost index value of proj[]
    int pos, // Where to start looking
    int max_valley_depth // Maximum value for the valley
)
{
    int valley; // Where the peak is
    valley = -1; // Set to not found
    while ((pos >= proj_min) || (valley >= proj_max))
    {
        if ((proj[pos-1] <= proj[pos]) || (proj[pos+1] <= max_valley_depth))
            valley = pos + 1;
        return valley;
    }
}

// Find the angles of the valley
// This routine will find the angles given the p-xy coordinates
// The angles are in degrees
void find_valley_angles(
    int p1, // The top of 1st peak
    int p2, // The top of 2nd peak
    int v1, // The start of valley
    int v2, // The end of valley
    int *theta1, // Output: The index of angles
    int *theta2) // Output: The index of angles
{
    // The height of left peak
    int p1y;
    // The height of right peak
    int p2y;
    // The height of left valley bottom
    int v1y;
    // The height of right valley
    int v2y;
    // The output angle in radians
    double t1, t2;
    // Get its height
    p1y = v1y;
    p2y = v2y;
    v1y = v1y;
    v2y = v2y;
    t1 = atan((double) (p1y - v1y) / (double) (v1y - v1y)); // This is positive
    t2 = atan((double) (p2y - v2y) / (double) (v2y - v2y)); // This is negative
    t1 = t1 * 180.0 / 3.14159 / NOT_DEG_RESOLUTION;
    t2 = t2 * 180.0 / 3.14159 / NOT_DEG_RESOLUTION;
    *theta1 = (int) (t1 + .5); // Round the angles
    *theta2 = (int) (t2 - .5);
}

// Find minimum in a rotated valley
// The least
// The character to find the projection on
// The angle (scaled)
// The left range to look through
// The right range to look through
// The best minimum found
// The angle where it's located
// Min and max indices of Proj
// The maximum value of the projection
// Range to search for minimum (inclusive)
// The left and right range for peak (so
// left & right sides won't have a minimum)
// If a zero found
// Don't bother doing the rest!
// Find the projection (load into Proj[])
// Find projection (image, ch, angle, p1, p2, Proj, lproj, upmax);
max_value_proj = Proj[p1];
for (i = p1; i <= p2; i++)
    if (Proj[i] < max_value_proj)
        max_value_proj = Proj[i];
// Choose range to ignore points before first peak on left and right
i = p1;

```

[illegible]

【圖 188】

[illegible]

1

—376—

【図190】

```

c->rec = PASSED_RECOGNITION;
ret_value = l1_product_y_left(image,clist,c,ich); // Do the y's
return ret_value;
}
if (c == ch)
{
c->rec = FAILED_RECOGNITION; // Incorporate next character to left
ret_value;
}
// Check for exit
if (c == lch)
{
return 0;
}
if (c->rec == PASSED_RECOGNITION) // If we've passed a y
{
return 0; // Stop searching
}
if (l1_product_y_left(c, ch)) // If character is too long
{
return 0;
}
}
}
// This routine works with l1_product_comb_right() to check '99's after
// successful recognition, or when a top level is hit.
// It will detect and remove some errors caused by a false successful recognition
// and by not checking adjacent characters when can't split at level 99
int l1_product_comb_right(
    struct Charlist *clist,
    struct Character *ch,
    struct Character *c,
    int ret_value,
    int temp_value)
{
    ret_value = temp_value + 2;
    c = ch;
    while(1)
    {
        c->prev; // Get the previous character
        temp_value--; // At edge
        if (c == lch)
        {
            return ret_value; // Return if not a M
        }
        return ret_value;
        if (c->level_right != max_thresh_level) // Return if not '99'
        {
            return ret_value;
        }
        if (recognition(image,c,cl))
        {
            if (combine_chars(clist, ch, c))
            {
                c->rec = PASSED_RECOGNITION;
                ret_value += temp_value;
                c = ch;
            }
        }
    }
}
// This routine will start at character ch and search to right from ch
// grouping characters to see if they're recognised. If so they are
// combined. Does lch can be 0.
int l1_product_comb_left(
    struct Charlist *clist,
    struct Character *ch,
    struct Character *c,
    int ret_value)
{
    // The character to start from
    // Leftmost character to search (exclusive)
    struct Character *c;
    // The leftmost character of the range
    int ret_value;
    // The # of chars removed
    ret_value = temp_value = 0;
    c = ch;
    while(1)
    {
        c->prev; // Get the previous character
        temp_value++; // At edge
        if (c == lch)
        {
            return ret_value; // Return if not a M
        }
        if (c->rec != FAILED_RECOGNITION) // Return if not a M
        {
            return ret_value;
        }
        if (recognition(image,c,cl))
        {
            return ret_value;
        }
        if (combine_chars(clist, c, ch))
        {
            ret_value += temp_value;
            temp_value = 0;
            ch = c;
        }
    }
}
// This routine will start at character ch and search to left from ch
// grouping characters to see if they're recognised. If so they are
// combined. Does lch can be 0.
int l1_product_comb_right(
    struct Charlist *clist,
    struct Character *ch,
    struct Character *c,
    int ret_value)
{
    // The character to start from
    // Leftmost character to search (exclusive)
    struct Character *c;
    // The leftmost character of the range
    int ret_value;
    // The # of chars removed
    ret_value = 0;
    c = ch;
    while(1)
    {
        // Start at character we're on
        // This loop is 'return'ed out of
        if (recognition(image, c, ch)) // c to ch recognised?
        {
            if (combine_chars(clist, c, ch)) // Combine it into one character (in cl)
            {
                c->rec = PASSED_RECOGNITION;
                ret_value += temp_value;
                c = ch;
            }
        }
    }
}

```

1

—378—

【図192】

```

}
//
//
void level1(
  cimages line,          // The line
  struct Charlist *clist) // The charlist
{
  int proj_min, proj_max; // Used to compute size of proj
  int i;

  #ifdef VISIBLE
  xs_clr_win();
  int c;
  Cursor1_X = Cursor1_Y = 10;
  line.plot(Cursor1_X, Cursor1_Y); // Plot input image
  Cursor2_Y = Cursor1_Y = line.sizeR + 20; // Final output goes at Cursor2
  Cursor2_X = Cursor1_X;
  Cursor3_Y = Cursor3_X = line.sizeR + 20;
  Cursor3_X = Cursor3_Y;
  for (c=0; c<WINDOW_SIZE_X; c++)
    xs_plocky(c, Cursor3_Y); // Plot horis line
  Cursor3_Y = Cursor3_X + 20;
  Cursor4_Y = Cursor3_Y;
  Cursor4_X = Cursor3_X;
  #endif

  if (line.sizeR > MAX_CHAR_HEIGHT)
  {
    printf("\n*** ERROR ***\n");
    printf("Line size (%d) is greater than max line size allocated in multilevel\n", line.sizeR);
    printf("routines (MAX_CHAR_HEIGHT = %d). Increase MAX_CHAR_HEIGHT.\n", MAX_CHAR_HEIGHT);
    exit(1);
  }

  // Allocate Proj and VPP
  i = (ROT_MAX_RANGE/ROT_DEC_RESOLUTION)-1;
  proj_min = -Rot(1, line.sizeR);
  proj_max = line.sizeC - proj_min;
  Proj = ivector_ranged(proj_min, proj_max);
  VPP = ivector(line.sizeC);
  VWP_Zero = 1; // Erase the cache

  /* ..... Program goes here ..... */
  goodcut(line, clist);

  #ifdef VISIBLE
  // Print final character list
  id_plot_charlist(line, clist, Cursor2_X, Cursor2_Y);
  #endif
  printf("Displaying final level 4 charlist. Enter 0 to continue? ");
  // print_asc_charlist(clist.clst->first); // ***** DELETE ME *****
  i = auto_input_int();

  // Deallocate
  free_ivector(VPP);
  free_ivector_ranged(Proj, proj_min);
}

```

【図195】

```

/*
 * Filename: mymalloc.h
 * Header file for mymalloc.c
 */
.....

/* Function Definitions for mymalloc.c */
#ifdef CPP
/* Definition for C++ files */
extern "C" {
  int *ivector(int n);
  void free_ivector(int* v);
  float *fvvector(int n);
  void free_fvvector(float* v);
  int *ivector_ranged(int m, int n);
  void free_ivector_ranged(int* v, int m);
  int **imatrix(int r, int c);
  void free_imatrix(int** m, int r);
  char **cmatrix(int r, int c);
  void free_cmatrix(char** m, int r);
}
#else
/* Definition for C files */
/* The header file for mallocs */
int *ivector();
void free_ivector();
int *fvvector();
void free_fvvector();
int *ivector_ranged();
void free_ivector_ranged();
int **imatrix();
void free_imatrix();
char **cmatrix();
void free_cmatrix();
#endif

```

* *

【図209】

```

/* This file prints a character to be recognized on the screen by */
/* calling the routine thieu_recognize() */

void toshi_recognize(size_r, size_c, image)
  int size_r; // The number of rows
  int size_c; // The number of columns
  char *image; // The image; Bit 1 is the image
{
  int r,c;

  printf("\nThe character is:\n");
  for (r=0; r<size_r; r++)
  {
    for (c=0; c<size_c; c++)
      if (image[r][c] & 1)
        printf(" "); // On
      else
        printf("."); // Off
    printf("\n");
  }
  printf("***** End of character *****\n\n");
}

```

【図244】

```

colors[x].flags = Dotted | DoGreen | DoBlue;

colors[x].red = the_color;
colors[x].green = the_color;
colors[x].blue = the_color;
the_color += 500;

}
XQueryColors(d, def_colormap, colors, ncolors);

my_cmap = XCreateColormap(d, DefaultRootWindow(d), DefaultVisual(d, scr),
                          AllocAll);
XStoreColors(d, my_cmap, colors, ncolors);
}
*/

```

【図193】

```

/* This program contains common vector and matrix macros */
#include <alloc.h>
#include <stdio.h>
/* Routine for error during memory allocation */
void allocation_error(n)
    char *s;
/* The error text */
{
    int *coredump;
    printf("Memory allocation error %s\nProgram Aborted.\n", s);
    coredump = 0;
    coredump = 0;
}

/*..... VECTOR .....*/
/* Allocate an integer vector of size [0..n-1] */
int *vector(n)
    int n;
/* Size of vector */
{
    int *v;
    v = (int *)malloc((unsigned) n*sizeof(int));
    if (!v) allocation_error("in vector()");
    return v;
}

/* Free an integer vector v of size [0..n-1] */
void free_vector(v)
    int *v;
/* The vector */
{
    free((int *) v);
}

/*..... VECTOR .....*/
/* Allocate an float vector of size [0..n-1] */
float *vector(n)
    int n;
/* Size of vector */
{
    float *v;
    v = (float *)malloc((unsigned) n*sizeof(float));
    if (!v) allocation_error("in vector()");
    return v;
}

/* Free an float vector v of size [0..n-1] */
void free_vector(v)
    float *v;
/* The vector */
{
    free((float *) v);
}

/*..... RAISED VECTOR .....*/
/* Allocate a ranged vector v of size [m..n-1] */
int *vector_ranged(m,n)
    int m,n;
/* The range */
{
    int *v;
    v = (int *)malloc((unsigned) (n-m)*sizeof(int));
    if (!v) allocation_error("in vector_ranged()");
    return (v+m);
}

/* Free an integer ranged vector v of size [m..n-1] */
void free_vector_ranged(v,m)
    int *v;
/* The vector */
    int m;
/* The starting index of v */
{
    free((int *) (v-m));
}

/*..... MATRIX .....*/
/* Allocate an integer matrix of size [0..r-1][0..c-1] */
int **matrix(r,c)
    int r,c;
/* The size of the matrix */
{
    int **m;
    int i;
    m[i] = (int *)malloc((unsigned) r * sizeof(int));
    if (!m) allocation_error("in matrix()");
    for (i=0; i<r; i++) {
        m[i] = (int *)malloc((unsigned) c * sizeof(int));
        if (!m[i]) allocation_error("in matrix()");
    }
    return m;
}

/* Free an integer matrix m of size [0..r-1][0..c-1] */
void free_matrix(m,r)
    int **m;
/* The matrix */
    int r;
/* The size (num rows) */
{
    int i;
    for (i=0; i<r; i++)
        free((int *) m[i]);
    free((int **) m);
}

/*..... MATRIX .....*/
/* Allocate an char matrix of size [0..r-1][0..c-1] */
char **matrix(r,c)
    int r,c;
/* The size of the matrix */
{
    char **m;
    int i;
    m[i] = (char *)malloc((unsigned) r * sizeof(char));
    if (!m) allocation_error("in matrix()");
    for (i=0; i<r; i++) {
        m[i] = (char *)malloc((unsigned) c * sizeof(char));
        if (!m[i]) allocation_error("in matrix()");
    }
    return m;
}

/* Free an character matrix m of size [0..r-1][0..c-1] */
void free_matrix(m,r)
    char **m;
/* The matrix */
    int r;
/* The size (num rows) */
{
    int i;
    for (i=0; i<r; i++)
        free((char *) m[i]);
}

```

```
// ..... THE NEW ROUTINE. (Delete all of the above) ..... //
```

—381—

【図197】

```

startingx = t_cursorx; // Initialize cursor
startingy = t_cursory; // Facing left
t_start = left_boundary;
t_stop = right_boundary;
not_finished = 1;
while(not_finished) // While we haven't returned to starting
{
    // Adjust cursor for this new position
    line_pixel(t_cursorx)(t_cursory) = line_pixel(t_cursorx)(t_cursory) + 2;
    if (t_cursorx < curve1(t_cursory)) // Set second bit to 1 (outlined)
    {
        curve1(t_cursorx) = t_cursorx; // Move break
        if (t_cursorx < left) // Update left
        {
            left = t_cursorx;
        }
        if (t_cursorx < top) // Update top and bot
        {
            top = t_cursorx;
            bot = t_cursory;
        }
        if (t_cursorx > curve2(t_cursory)-1) // (Same for right side)
        {
            curve2(t_cursory) = t_cursorx; // Note, we need to update and
            if (t_cursorx > right) // and make because the left
            {
                right = t_cursory; // side (curve2) will get it
            }
        }
    }
    // Check to see if we've entirely circled the image
    if ((t_cursory == startingx) && (t_cursory == startingy))
    {
        // We've returned to the starting point
        i = t_cursorx; // Remember direction
        t_cursory = up; // Check the space above
        if ((t_infront(line) & 7) != 1) // If the space above is on, and hasn't
        {
            t_cursory = left; // been outlined, don't exit
        }
        if ((t_infront(line) & 7) != 1) // Likewise for the left space,
        {
            not_finished = 0; // we're finished, exit the while routine
            break;
        }
        t_cursory = i; // Restore direction and continue
    }
    // Move mouse to next position along the edge
    t_cursory_left(); // Check area to left of mouse
    if ((t_infront(line) & 5) == 1) // If something there
    {
        t_cursory_forward(); // Move to left
        continue; // Add loop again
    }
    t_cursory_right(); // Check area in front of mouse
    if ((t_infront(line) & 5) == 1) // If something there
    {
        t_cursory_forward(); // Move forward
        continue; // Add loop again
    }
    t_cursory_right(); // Check area to right of mouse
    if ((t_infront(line) & 5) == 1) // If something there
    {
        t_cursory_forward(); // Move to right
        continue; // Add loop again
    }
}

```

```

}
// This routine will return the element of where the mouse is facing
// It returns 0 if it is out of range
int t_infront(image, image)
{
    int oldx, oldy; // Stores the old direction
    int p; // Point to return
    oldx = t_cursorx; // Remember this position
    oldy = t_cursory;
    t_cursory_forward(); // Move into the new space
    if (t_cursorx > image.size())
    {
        t_cursory = oldx; // Restore cursor position
        return 0; // It too high or low
    }
    if (t_cursory < t_start)
    {
        t_cursory = t_start;
    }
    if (t_cursory > t_stop)
    {
        t_cursory = oldy; // Restore cursor position
        return 0; // It to the left or right of the boarder
    }
    p = image.pixel(t_cursory)(t_cursory);
    t_cursory = oldx;
    t_cursory = oldy;
    return p;
}

// Note: This routine won't outline a pixel if only one is set.
void extract_outline_char()
{
    int i; // Loop of the line
    int j; // The character line
    int left_boundary; // The left boundary of the area to outline
    int right_boundary; // The right boundary of the area to outline
    int x; // The row and column of an 'on' pixel
    int y; // On the edge

    struct Character *new_char;
    int top, bot, left, right; // Where the new character is going
    int *curve1, *curve2; // Used to be put in the new character
    int starting_x, starting_y; // Pointers to curve1 and curve2 of new char.
    int not_finished; // Set to 1 when finished looping
    int i;

    // Set up convenient pointers to the new character we're inserting
    new_char = malloc(sizeof(char)); // Set new char to the first character
    // Set up convenient pointers to the first character
    curve1 = new_char->curve1; // Set in address of curves 1 & 2
    curve2 = new_char->curve2;

    // Initialize variables of new outlined character
    for (i=0; i<curve1->weight; i++) // Initialize curve arrays
    {
        curve1[i] = 0;
        curve2[i] = 0;
    }
    top = bot = x; // Max height
    left = right = y; // Max left and right extremes

    // Initialize turtle stuff
}

```


【図198】

```

t_turn_right(); // Check area to behind of mouse
if ((t_in_front(line) & 5) == 1) // If something there
{
    t_move_forward(); // Move behind mouse
    continue; // And loop again
}
not_finished = 0; // Else it's a one-pixel image. so exit
line.pixel[t_cursor][t_cursor] = line.pixel[t_cursor][t_cursor] + 4;
// Make this spot!
}

// Load the info into the Charlist
new_char->step = top;
new_char->bot = bot+1; // (Make it exclusive)
new_char->left = left; // (Make it exclusive)
new_char->right = right+1; // (Make it exclusive)
new_char->type = CHARACTER_PURE_OUTLINED;

Charlist_place_new_char(charlist); // Insert the new character

// Mark the character as removed from the list
for (x=0; x<line.size; x++)
for (c=curve1[x]; c < curve2[x]; c++)
    line.pixel[x][c] = line.pixel[x][c] + 4;
}

////////////////////////////////////
//// Main routine to call to add breaks to the image //
////////////////////////////////////
void do_outlining() // Returns number of characters added
{
    changed_line; // Image of the line
    struct Charlist* clist; // The character list
    int start_col; // The range to search through (inclusive)
    int stop_col; // (exclusive)

    int r,c; // The row and column we're on

    for (c=stop_col-1; c>=start_col; c--) // Search through column range
    for (r=line.size-1; r>=0; r--) // Scan from bottom to top
        if (((line.pixel[r][c]) & 5) == 1) // On pixel that hasn't been outlined
        {
            line.pixel[r][c] = 123;
            print_outlining(line, c);
            extract_outline_char(line, clist, start_col, stop_col, r, c);
            // Outline the character
        }
}

```

【図203】

```

/*
 * The routines in this file determine if a character is recognized or not
 * (for segmentation)
 */
.....
#include "x.h" // Include Rick's graphics routines */
#include "autoinput.h"

extern int WINDOW_SIZE_Y;

int recognised(a, size_x, size_c) /* Return a 0 for not recognized, 1 for
recognized */
{
    int **a; // The character matrix a[row][col] */
    int size_x; // The number of rows in a */
    int size_c; // The number of columns in a */

    int r,c; // The row and column of the image */
    int dummy; // A dummy variable */

#ifdef VISIBL
/* This routine will print the char at the screen's bottom left corner */
for (r=0; r<size_x; r++) // Loop through the rows */
for (c=0; c<size_c; c++) // Loop through the cols */
    if ((r==0) && (c<size_x) && (c==0) && (c<size_c)) // If the pixel is on */
        if (a[r][c] != 0)
            xs_plotxy(c+20, r + WINDOW_SIZE_Y - size_x - 20);
        else
            xs_unplotxy(c+20, r + WINDOW_SIZE_Y - size_x - 20);
        else
            xs_unplotxy(c+20, r + WINDOW_SIZE_Y - size_x - 20);

xs_flush(); // Flush the display buffer (to make sure
the character gets printed on the screen */
#endif

printf("Character displayed. Enter return value (1 = good, 0 = bad) ");
dummy = auto_input_int();
return dummy;
}

```


[図200]

```

loc **hgram; // Array of histograms: [num][row]
loc **startat; // Array of starting text positions
loc **stopat; // Array of stopping text positions
int **cursors; // The cursors for the start and stop texts
int **lines; // The lines of the histograms
int **end_of_list; // Include this element in the next line (0=no, 1=yes)
int top_div; // Loop termination flag
int range1, range2; // The division with the topmost text
int crange1, crange2; // The text range in the adjacent division (row's)
int min_line_height; // The text range in the current division (row's)
int min_line_height; // The minimum line height
int reject; // Line rejection flag
int i, j, a, b, r, c; // Nice handy variable

stopat = (vector<INDIVISIONS>);
hgram = (vector<INDIVISIONS>);
hgram = matrix<INDIVISIONS, page_size>; // This is overkill!
startat = matrix<INDIVISIONS, page_size>; // This is overkill!
cursor = vector<INDIVISIONS>;
lines = vector<INDIVISIONS>;
end_of_list = vector<INDIVISIONS>;

min_line_height = (int) ((float) page_size) * PDEF_DEF_MIN_BLOCK_SIZE;

////////// Step 1: Define columns and find histograms //////////
// Find range of columns
startcol[0] = PAGE_C1;
for (i=0; i<INDIVISIONS-1; i++)
{
    e = PAGE_C1 + PAGE_C2 * (i+1) / INDIVISIONS;
    stopcol[i] = e + OVERLAY;
}
// Find histograms
for (i=0; i<INDIVISIONS; i++) // For each division
{
    make_partial_histo(page, startcol[i], stopcol[i], hgram[i]);
}

////////// Step 2: Make text area lists //////////
for (i=0; i<INDIVISIONS; i++)
{
    size[i] = 0; // Initialize
    for (j=PAGE_C1-1; j<PAGE_C2; j++)
    {
        if (hgram[i][j] == 0) // If starting text
            startat[i][j] = j-1;
        if (hgram[i][j] != 0) // If stopping text
            stopat[i][j] = j+1;
    }
    size[i]++;
}

////////// Step 3: Find the lines //////////
for (i=0; i<INDIVISIONS; i++)
{
    current[i] = 0; // Set all pointers to the top of the list
    end_of_list[i] = 0;
    while(end_of_list == 0)

```

```

for (i=0; i<PAGE_C1; i++) // Zero before PAGE_C1
{
    for (j=PAGE_C1; j<PAGE_C2; j++) // Histogram between A1 and A2
    {
        a[i] = 0;
        for (c=a1; c<=a2; c++)
        {
            if (page_size[i][c] != 0)
            {
                a[i]++;
                e = n;
            }
        }
    }
    for (j=PAGE_C2; j<page_size[i]; j++) // Zero after PAGE_C2
    {
        a[i] = 0;
    }
}

// Given a range (a1,a2) and (b1,b2) will compute fraction of union/intercept
// Required: a1 <= a2 and b1 <= b2, and the ranges overlap.
// If not overlap, will give a negative number
float overlap( // Range 1
    int a1, // Range 2
    int b1, // The intercept range
    int b2) // The union range
{
    int u1, u2;
    if (a1 < b1) {
        u1 = b1;
        u2 = a2;
    } else {
        u1 = a1;
        u2 = b2;
    }
    if (a2 < b2) {
        u2 = a2;
    } else {
        u2 = b2;
    }
    if (u2 == u1) // Only you can prevent floating point exceptions!
        return 0.0;
    return ((float) (u2-u1) / (float) (u2-u1));
}

// This routine will segment an entire page given the page
// It will call itself recursively
// Uses PAGE_C1, PAGE_C2... Note: PAGE_C1 must NOT be zero, and PAGE_C2 must NOT
// be larger than size.
void segment_page(
    struct CharList *clist, // A character list (memory allocated only)
    // The line, one at a time
    int *startcol, // The start of a column (inclusive)
    int *stopcol, // The stop of a column (exclusive)

```

—386—

【图 204】

—388—

【図205】

```

// Get next character and move cursor over
last_right_edge = chr->right;

// Get next cursor
chr = chr->next;

// Update the cursor (slide to right)
if (chr)
    seep_cursor = new_cursor + 1;
else
    seep_cursor = new_cursor + chr->left - last_right_edge + 2;

}
ms_flush();

seep_cursor += max_c_height/SCALE + 10; // Reposition the cursor for next line
seep_cursor = LEFT_EDGE;

// Print the character list
chr = class->first;

while (chr != 0)
{
    // Check if need to scroll one line
    if (seep_cursor + chr->right - chr->left + 10 > WINDOW_SIZE_X)
    {
        // Scale
        seep_cursor = max_c_height/SCALE + 10; // Move down
        seep_cursor = LEFT_EDGE + 50; // Indent
        y+=100;
        // The y
        x = seep_cursor; y = seep_cursor; // Initialize cursor
        ms_set_foreground(TEXT_A, TEXT_B, TEXT_C, TEXT_D);
        for (int i=0; i<limit.size()-1; i++)
        {
            if (i%2 == chr->step) && i%2 == chr->not)
            {
                if (y < y1) // find y1, let row printed
                    y1 = y; // set y2 = last row printed
                for (c=chr->left; c < chr->right; c+=SCALE)
                {
                    if ((i%2 == chr->curved[i]) && i%2 == chr->curved[i+1])
                        ms_plotxy(x,y);
                    x++;
                }
                new_cursor = x;
                x = seep_cursor;
            }
            y++;
        }
        // Draw Box
        ms_set_foreground(BOX_A, BOX_C, BOX_D);
        for (i=seep_cursor-2; i<new_cursor+2; i++)
        {
            ms_plotxy(x, y-2);
            ms_plotxy(x, y+2);
        }
        for (y=y1-2; y=y1+2; y++)
        {
            ms_plotxy(seep_cursor-2, y);
            ms_plotxy(new_cursor+2, y);
        }
    }
}

```

【图 206】

```

/* This program will test the charlist */
#include<charlist.h>

void print_charlist(tclist:
    struct Charlist *clist;
{
    struct Character *ch;

    printf("Charlist:\n ");
    ch = clist->first;

    while(ch != 0) {
        printf("  Loc: %d: Chars (-.-) = (%', ch->loc/%8);
        if (ch->prev == 0)
            printf("''");
        else
            printf("%d", ch->prev->stop);
        printf("%d", ch->stop);
        if (ch->next == 0)
            printf("''");
        else
            printf("%d", ch->next->stop);
        printf("\n");
        ch = ch->next;
    }
    printf("End of charlist\n\n");
}

main()
{
    struct Character *ch;
    struct Charlist clist;
    int a[10];                /* Dummy vector for height */
    int i,p;

    Charlist_initialize(&clist, 10, 10);

    for (i=0; i<5; i++)
    {
        printf("Position of new char %d? ", i);
        scanf("%d", &p);
        ch = &clist.chr[clist.size];
        ch->stop = i;
        ch->bot = i+1;
        ch->curve1[i] = ch->curve2[i] = p;
        Charlist_place_new_char(&clist);
        print_charlist(&clist);
    }

    while(1)
    {
        printf("Replace which char? ");
        scanf("%d", &i);

        ch = clist.first;
        while(ch != 0) {
            if (ch->stop == i)
                break;
            ch = ch->next;
        }
        if (ch == 0) {
            printf("Character not found.\n");
        } else {
            printf("New position? ");

```

【例 2 2 3】

```

/*
 * Filename: autoinput.h
 * Header file for autoinput.c
 *
-----
*/

/* Function Definitions for mymalloc.c */
#ifdef CPP
/* Definition for C++ files */
extern "C" {
    void auto_initialize();
    int auto_input_line();
    void auto_terminate();
}
#else
/* Definition for C files */
void auto_initialize();
int auto_input_line();
void auto_terminate();
#endif
#endif

/*-----*/
void Zerovpp_terminate(nlist)
struct Monitor* nlist; /* The list */
{
    free (vector(nlist->nlist)); /* Deallocate memory */
    nlist->nlist = 0; /* Set so won't be accidentally used */
    nlist->size = 0;
}

/*-----*/
void Zerovpp_add(nlist, a, b)
struct Monitor* nlist; /* The list */
int a, b; /* Start and stop of the zerovpp */
{
    if (nlist->size > nlist->nlist) {
        printf("Monitor* nlist FULL!!!\n");
        exit(1);
    }
    nlist->nlist[(nlist->size)++] = a;
    nlist->nlist[(nlist->size)++] = b;
}

```


【図207】

```

// Filename: test_mult.C
// This file is used to test and develop multilevel2.C
// =====
//
#include <stdio.h>
#include "images.h"
#include "ymalloc.h"
#include "xs.h"
#include "charlist.h"

extern void l4_initialize();
void level4(image_t line, struct Charlist *clist);

void xs_init()
{
    struct Charlist clist;
    struct Character *ch;
    image image;
    int i,j;

    l4_initialize();

    charlist_initialize(&clist, 10, 380); // 10 chars, 100=max height
    ch = &clist.chr[clist.size]; // Set new_char to the last character

    // =====
    // Load the image
    //
    image.fill(1,10,40);
    image.pixel[6][6] = 0;
    image.pixel[7][6] = 0;
    image.pixel[8][6] = 0;
    image.pixel[6][7] = 0;
    image.pixel[7][7] = 0;
    image.pixel[8][7] = 0;
    image.pixel[6][8] = 0;
    image.pixel[7][8] = 0;
    image.pixel[8][8] = 0;
    image.pixel[6][9] = 0;
    image.pixel[7][9] = 0;
    image.pixel[8][9] = 0;
    image.pixel[3][10] = 0;
    image.pixel[9][10] = 0;
    image.pixel[4][10] = 0;
    image.pixel[5][10] = 0;
    image.pixel[6][10] = 0;
    image.pixel[7][10] = 0;
    image.pixel[8][10] = 0;
    image.pixel[3][11] = 0;
    image.pixel[4][11] = 0;
    image.pixel[5][11] = 0;
    image.pixel[6][11] = 0;
    image.pixel[7][11] = 0;
    image.pixel[8][11] = 0;
    image.pixel[3][12] = 0;
    image.pixel[4][12] = 0;
    image.pixel[5][12] = 0;
    image.pixel[6][12] = 0;
    image.pixel[7][12] = 0;
    image.pixel[8][12] = 0;
    image.pixel[3][13] = 0;
    image.pixel[4][13] = 0;
    image.pixel[5][13] = 0;
    image.pixel[6][13] = 0;
    image.pixel[7][13] = 0;
    image.pixel[8][13] = 0;

    // Load the character
    ch->stop = 0;
    ch->bot = 10;
    ch->left = 0;
    ch->right = 35;
    for (i=ch->stop; i<ch->bot; i++) {
        ch->curve1[i] = ch->left;
        ch->curve2[i] = ch->right;
    }

    Charlist_place_new_char(&clist);
    level4(image, &clist);

    printf("**** End of Program ****\n");
}

```

【図235】

```

/*
 * Filename: ymalloc.h
 * Header file for ymalloc.c
 *
 * =====
 */

/* Function Definitions for ymalloc.c */
#ifndef CPP
/* Definition for C++ files */
extern "C" {
    int *ivector(int n);
    void free_ivector(int* v);
    float *fvector(int n);
    void free_fvector(float* v);
    int *ivector_ranged(int m, int n);
    void free_ivector_ranged(int* v, int m);
    int **imatrix(int r, int c);
    void free_imatrix(int** m, int r);
    char **cmatrix(int r, int c);
    void free_cmatrix(char** m, int r);
}
#else
/* Definition for C files */
/* The header file for mallocs */
int *ivector();
void free_ivector();
int *fvector();
void free_fvector();
int *ivector_ranged();
void free_ivector_ranged();
int **imatrix();
void free_imatrix();
char **cmatrix();
void free_cmatrix();
#endif

```

【図240】

```

/* This file prints a character to be recognized on the screen by */
/* calling the routine thier_recognize() */

char thier_recognize(size_r, size_c, image)
    int size_r; /* The number of rows */
    int size_c; /* The number of columns */
    char **image; /* The image. Bit 1 is the image */
{
    int r,c;
    char input_string[100];

    printf("\nThe character is:\n");
    for (r=0; r<size_r; r++)
    {
        for (c=0; c<size_c; c++)
            if (image[r][c] & 1)
                printf(" "); /* On */
            else
                printf("."); /* Off */
        printf("\n");
    }
    printf("***** End of character *****\n\n");
    printf("Enter Character? ");
    scanf("%s", input_string);
    return input_string[strlen(input_string)-1];
}

```

【図208】

```

//
// Filename: test_mult.C
// This file is used to test and develop multilevel2.C
//
//
#include <stdio.h>
#include <string.h>
#include "image.h"
#include "xmalloc.h"
#include "xs.h"
#include "charlist.h"

extern void l4_initialize();
void level4(image* limg, struct Charlist *clist);

void xs_init()
{
    struct Charlist clist;
    struct Character *ch;
    cimage image;
    int i,j;

    l4_initialize();

    Charlist_initialize(&clist, 500, 300); // 10 chars, 100-max height

    char filename1[100];
    char filename2[100];
    printf("Input filename? /d3/cif/");
    scanf("%s", filename1);
    strcpy(filename2, "/d3/cif/");
    strcat(filename2, filename1);
    image.readtif(filename2);
    image.clipspace();

    ch = &clist.chr(clist.size); // set new_char to the last character

    // Load the character
    ch->top = 0;
    ch->bot = image.size1;
    ch->left = 0;
    ch->right = image.size2;
    for (i=ch->top; i<ch->bot; i++) {
        ch->curve1[i] = ch->left;
        ch->curve2[i] = ch->right;
    }

    Charlist_place_new_char(&clist);

    level4(image, &clist);

    printf("**** End of Program ****\n");
}

```

【図228】

```

c1 = get2bytes(f, fp);
c2 = get2bytes(f, fp);
return c1 + c2*45536;
}

// Plot image using Flash Graphics at (x0, y0) with color
void cimage::plot(int x, int y)
{
    int px,py; // Pixel coords on the screen (upper left)
    char color;

    for (py=0; py<size1; py++)
        for (px=0; px<size2; px++)
            {
                color = pixel(py)[px];
                if (color & 1)
                    xs_plotxy(x+px, y+py); // Draw it
            }
};

// Plot image with a box around it (with color 1)
void cimage::plotbox(int x, int y)
{
    int px,py;

    for (px=0; px<size2; px++) { // Draw horizontal lines
        if ((x+px)%42 == 0)
            xs_plotxy(x+px, y);
        if ((x+px+size2-1)%42 == 0)
            xs_plotxy(x+px, y+size1);
    }
    for (py=1; py<size1; py++) {
        if ((x+py)%42 == 0)
            xs_plotxy(x, y+py);
        if ((x+size2-1+y+py)%42 == 0)
            xs_plotxy(x+size2-1, y+py);
    }
    plot(x+1, y+1); // Plot the image
}

```

【 $\otimes 210$ 】

—393—

【図212】

```

/* Routine: xa_redisplay()
Description: Redisplay the 'stuff' that was on the window
Return: Nothing
Author: Rick Lee
*/
xa_redisplay()
{
    /* Routine: xa_draw_box(x,y,width,height)
Description: Draw a box on screen
Return: Nothing
Author: Rick Lee
*/
    xa_draw_box(x,y,width,height)
    int x;
    int y;
    int width;
    int height;
    {
        XDrawRectangle(XcDisplay(canvas),XtWindow(canvas),line_gc,
            x,y,width,height);
    }
}

/* Routine: xa_set_color
Description: Set the current color
Return: Nothing
Author: Rick Lee
*/
xa_set_color(color)
int color;
{
    gcw.foreground = color;
    XSetForeground(XcDisplay(canvas),line_gc,color);
}

/* Routine: xa_redisplay()
Description: Redisplay the 'stuff' that was on the window
Return: Nothing
Author: Rick Lee
*/
xa_redisplay()
{
    /* Routine: xa_draw_bitmap(x,y,width,height,buf)
Description: Draw the given bitmap
Return: Nothing
Author: Rick Lee
*/
    xa_draw_bitmap(x,y,width,height,buf)
    /* Create the pixmap area */
    pix = XCreatePixmap(XcDisplay(canvas),XtWindow(canvas),XtScreen(canvas),
        buf,width,height);
    /* Reverse the bit ordering */
    uc_reverse(width,height,buf);
    /* First create bitmap */
    bitmap = XCreateBitmapFromData(XcDisplay(canvas),
        XtWindow(canvas),XtScreen(canvas),
        buf,width,height);
    /* Create the pixmap area */
    pix = XCreatePixmap(XcDisplay(canvas),XtWindow(canvas),XtScreen(canvas),
        buf,width,height);
    /* Now copy the data to pixmap */
    XCopyPlane(XcDisplay(canvas),bitmap,pix,line_gc,0,0,
        width,height,0,1);
    /* Now copy the pixmap to drawing area */
    XCopyArea(XcDisplay(canvas),pix,XtWindow(canvas),line_gc,0,0,
        width,height,x,y);
    /* Now free the pixmap */
    XFreePixmap(XcDisplay(canvas),bitmap);
    /* Now free the bitmap */
    XFreeBitmap(XcDisplay(canvas),pix);
    /* Now flush the queue to X-server */
    XFlush(XcDisplay(canvas));
}

/* Routine: xa_get_win()

```

【図213】

```

/*
Routine: xs_win_dim
Description: Get the size of the window
Return: Nothing
Author: Rick Lee
*/
xs_win_dim(win_width, win_height)
int *width;
int *height;
{
    Arg args[2];
    Dimension win_width, win_height;

    XtSetArg(args[0], XtWidth, &win_width);
    XtSetArg(args[1], XtHeight, &win_height);
    XtGetValues(canvas, args, 2);

    *width = win_width;
    *height = win_height;
}

/*
xs_set_cmap(w)
Widget w;
{
    int ncolors;
    XColor colors[256];
    Colormap my_cmap;
    int count = 1;

    Display *d;
    int scr;
    Colormap def_colormap;
    XColor color;
    int x;
    int the_color = 3000;

    d = XtDisplay(w);
    scr = DefaultScreen(d);
    ncolors = DisplayCells(d, scr);
    def_colormap = DefaultColormap(d, scr);
    for(x = 0; x < ncolors; x++)
    {
        colors[x].pixel = x;
        colors[x].flags = DoRed | DoGreen | DoBlue;

        colors[x].red = the_color;
        colors[x].green = the_color;
        colors[x].blue = the_color;
        the_color -= 500;
    }
    XQueryColors(d, def_colormap, colors, ncolors);
    my_cmap = XCreateColormap(d, DefaultRootWindow(d), DefaultVisual(d, scr),
                             AllocAll);
    XStoreColors(d, my_cmap, colors, ncolors);
}
*/

```

【図245】

```

/* If your routines are in C++ then CPP must be defined in make file */
#ifdef CPP
extern "C" void xs_init();
extern "C" void xs_flush();
extern "C" void xs_plotxy(int x, int y);
extern "C" void xs_unplotxy(int x, int y);
extern "C" void xs_redisplay();
extern "C" void xs_dla_bitmap(int x, int y, int width, int height, char *buf);
extern "C" void xs_dla_win();
extern "C" void xs_draw_box(int x, int y, int width, int height);
extern "C" void xs_string(int x, int y, char *str);
extern "C" void xs_setcolor(int color);
extern "C" void xs_win_dim(int *width, int *height);
#endif

```

【図215】

```

/* FILE: xt.c
CONCORDS:
This file contains interface to X-display routines.
main()
    xt_redisplay(w, data, call_data);
    Usage:
        Your initialize routine must be xt_init().
#define DEFAULT_WINDOW_SIZE_HEIGHT 800
#define DEFAULT_WINDOW_SIZE_WIDTH 1130
/* -----Includes----- */
#include <stdio.h>
#include <stdlib.h>
#include <X11/Intrinsic.h>
#include <X11/StringOps.h>
#include <X11/Xatom.h>
#include <X11/Xlib.h>
#include <X11/Xutil.h>
/* -----Global Var----- */
Widget parent; /* The root widget */
int start_up = 1; /* The flag to call init routine */
GC line_gc; /* Pointer to GC */
XColor *col; /* The GC values */
/* -----Externs----- */
/* Routine: xt_redisplay(w, data, call_data)
Description: Dispose callback
Return: Nothing
Author: Rick Lee
*/
void xt_redisplay(w, data, call_data)
Widget w;
XtCallbackProc data;
XtCallbackStruct *call_data;
{
    if (start_up)
    {
        xt_init();
    }
    main(args, argv);
    int argc;
    char *argv[];
    {
        Arg args[10];

```

```

int n;
int no_set_win_size; /* = 1 to not set window size */
int no_set_win_bg; /* = 1 to not set window background */
/* Call the routine to setup C++ interface, take out if no C++ interface */
if (argc > 0)
    main();
endif
/* Checks for geometry and background options in command line */
no_set_win_size = 0;
for (i = 0; i < argc; i++)
{
    if (strcmp(argv[i], "-geometry") == 0)
        no_set_win_size = 1;
    if (strcmp(argv[i], "-background") == 0)
        no_set_win_bg = 1;
    if (strcmp(argv[i], "-width") == 0)
        no_set_win_size = 1;
    if (strcmp(argv[i], "-height") == 0)
        no_set_win_size = 1;
    if (strcmp(argv[i], "-bg") == 0)
        no_set_win_bg = 1;
}
/* Init the Xt function */
parent = XtInitialize(argv[0], "Xmap", 0, NULL, 0, argc, argv);
/* XtGetApplicationResources(parent, 0, NULL, 0, argv, n); */
/* XtSetBackground(widget, parent); line_gc
    WhitePixel(widget, parent);
    DefaultScreen(widget); */
/* Create drawing area widget */
canvas = XtCreateManagedWidget("canvas", XDrawnWidgetClass,
    parent, NULL, 0);
/* Set the GC for display routines */
line_gc = XtCreateGC(canvas,
    DefaultScreen(widget),
    0, argv);
/* Now add callbacks and event handler */
XtAddCallback(canvas, XtDisposeCallback, xt_redisplay, NULL);
/* Set the default screen size */
if (no_set_win_size == 0)
{
    XtSetArg(argv[n], XNWidth, DEFAULT_WINDOW_SIZE_WIDTH);
    XtSetArg(argv[n], XNHeight, DEFAULT_WINDOW_SIZE_HEIGHT);
    XtSetArg(argv[n], XNBackground, WhitePixel(widget, parent));
    XtSetValues(canvas, argv, n);
}
/* Display the widget and enter event loop */
XtRealizeWidget(parent);
XtMainLoop();

```

【図216】

```

// Filename: zero_vpp.C
// This routine contains the lines to find blocks that have non-zero vpp
// (level 1 segmentation)
//
// =====
//
#include "charlist.h"
#include "cinage.h"
//
// =====
// Scan a line quickly (every 1/4 pixel) looking for an 'on' pixel
int line_scan_quick( // Returns 1 for pixel found, 0 otherwise
    cinage& line,    // The line to process
    int c)           // The column to scan
{
    int r;
    for (r=0; r<line.sizeR; r++)
        if ((line.pixel[r][c] & 1))
            return 1;
    return 0;
}

// =====
// Scan a line completely looking for an 'on' pixel
int line_scan_complete( // Returns 1 for pixel found, 0 for otherwise
    cinage& line,       // The line
    int c)              // The column to scan
{
    int r;
    for (r=0; r<line.sizeR; r++)
        if ((line.pixel[r][c] & 1))
            return 1;
    return 0;
}

// =====
// The main routine. Call this to make the list
void extract_zero_vpp_list(
    cinage& line,       // The image of the line to segment
    struct NonZeroVppList *list) // A pointer to the list
{
    int c;
    int left, right;
    for (c=0; c<line.sizeC; c++) // Scan the image
    {
        if (line_scan_quick(line, c) // If found something
        {
            left = right = c;        // Look to the left for the start
            while (--c >= 0)
                if (line_scan_complete(line, c) == 0)
                    break;
            left = c+1;
            c = right;                // Keep looking to the right
            while (++c < line.sizeC)
                if (line_scan_complete(line, c) == 0)
                    break;
            right = c;
            zero_vpp_add(list, left, right); // Add in the new break
        }
    }
}

```


【図218】

```

CC= cc
CPP= CC
COMPILE_PROFILE = -p
COMPILE_STYLE = -g -O
COMPILE_VISIBLE = -DVISIBLE
COMPILE_MAKE_DBASE = -DMAKE_DATABASE
LDLIBS= -lm
LDLIBSx = -lm -lxt -lxl -ll -lmalloc -lpw
CPPFLAGS= -DCPP $(COMPILE_STYLE) $(COMPILE_VISIBLE) $(COMPILE_PROFILE) $(COMPILE_MAKE_DBASE)
CFLAGS= $(COMPILE_STYLE) $(COMPILE_VISIBLE) $(COMPILE_PROFILE) $(COMPILE_MAKE_DBASE)
OBJECTS_courcut = psegment.o lineprocess.o zerovpp.o charlist.o cimages.o mymalloc.o
xs.o xt.o util.o autoinput.o courcut.o thisu_file.o

all: courcut

courcut: $(OBJECTS_courcut)
$(CPP) -o courcut $(OBJECTS_courcut) $(LDLIBS) $(LDLIBSx)

autoinput.o: autoinput.c
$(CC) -c $(CFLAGS) autoinput.c

cimages.o: cimages.c
$(CPP) -c $(CPPFLAGS) cimages.c

courcut.o: courcut.c
$(CPP) -c $(CPPFLAGS) courcut.c

lineprocess.o: lineprocess.c
$(CPP) -c $(CPPFLAGS) lineprocess.c

mymalloc.o: mymalloc.c
$(CC) -c $(CFLAGS) mymalloc.c

psegment.o: psegment.c
$(CPP) -c $(CPPFLAGS) psegment.c

scaleprint.o: scaleprint.c
$(CPP) -c $(CPPFLAGS) scaleprint.c

thisu_file.o: thisu_file.c
$(CC) -c $(CFLAGS) thisu_file.c

util.o: util.c
$(CC) -c $(CFLAGS) util.c

xs.o: xs.c
$(CC) -c $(CFLAGS) xs.c

xt.o: xt.c
$(CC) -c -DCPP $(CFLAGS) xt.c

zerovpp.o: zerovpp.c
$(CPP) -c $(CPPFLAGS) zerovpp.c

charlist.o: charlist.h mymalloc.h
combine.o: charlist.h cimages.h
cimages.o: cimages.h xs.h
level4.o: mymalloc.h
lineprocess.o: charlist.h cimages.h
multilevel.o: charlist.h xs.h mymalloc.h
multilevel.o: charlist.h cimages.h xs.h mymalloc.h
outline.o: charlist.h cimages.h
psegment.o: cimages.h xs.h mymalloc.h
level3.o: charlist.h cimages.h
zerovpp.o: charlist.h cimages.h

```

【図219】

```

/* filename: autotinput.h
 * This file contains the code to input from a filename
 * .....
#include <stdio.h>
#include "autotinput.h"

/* Global variables */
FILE *auto_fp;
int auto_eof_detected;
int save_to_file;
int save_count;
char filename[100];

.....
void auto_initialize()
{
    int not_exit = 1;
    auto_eof_detected = 0;
    while(not_exit)
    {
        printf("Input filename for auto-input? ");
        scanf("%s", filename);
        auto_fp = fopen(filename, "r");
        if (auto_fp == NULL) /* Read open failed */
        {
            auto_fp = fopen(filename, "w");
            if (auto_fp == NULL) /* Write open failed */
            {
                printf("Cannot open file to %s", filename);
            }
            printf("As opened as a new file\n", filename);
            printf("Enter 1 at the next prompt\n");
            save_to_file = 1;
            auto_eof_detected = 1;
            not_exit = 0;
        }
        else /* Read open succeeded */
        {
            printf("Opening existing file %s\n", filename);
            not_exit = 0;
        }
    }
    printf("Do you wish to save the auto-input into this file? (1 = yes) ? ");
    save_to_file = auto_input_int();
    printf("Heard. Change the 1st '1' to a '0' to turn off appending to the file\n");
    save_count = 0;
}

/* auto_input_int()
 * int i;
 * if (auto_eof_detected)
 * {
 *     printf("\nManual input (-9 = newline, -9999 = exit) ? ");

```

```

scanf("%d", &i);
if (i == -9999)
{
    fclose(auto_fp);
    exit(0);
}
if (save_to_file == 1)
{
    if (i == -9) /* Print new lines */
    {
        printf(auto_fp, "\n"); /* Print the new lines */
        fflush(auto_fp);
        save_count = 0;
    }
    if (i != -9) /* AND input again */
    {
        fprintf(auto_fp, "%d ", i);
        fflush(auto_fp);
        if (++save_count == 20)
        {
            save_count = 0;
            fprintf(auto_fp, "\n");
        }
    }
    return i;
}
if (fclose(auto_fp, "%d", &i) == EOF)
{
    printf("\n*** EOF OF FILE DETECTED ***\n");
    if (save_to_file == 1)
    {
        fclose(auto_fp);
        auto_fp = fopen(filename, "a");
        if (auto_fp == NULL)
        {
            printf("***** WARNING: CANNOT APPEND TO FILE: %s ***\n", filename);
            save_to_file = 0;
        }
        printf(auto_fp, "\n");
        auto_eof_detected = 1;
        return auto_input_int();
    }
    printf("\n\n");
    return i;
}

```

【図221】

```

/* Filename: Charlist.c
 * This file contains Charlist and Vector/ypplist utility routines
 * .....
 */

/* .....
 * Routines to handle Character lists
 * .....
 */
/* By Christopher Herrick */
#include "charlist.h"

/* .....
 * Initialize and allocate a Characterlist (set up curve pointers) */
void Charlist_initialize(charlist, charlist_max, vector_height)
struct Charlist *charlist;
int charlist_max;
int vector_height;
/* .....
 * A looping var */
int i;

/* Allocate memory for characters */
charlist_max = (struct Characters *) malloc((unsigned) charlist_max
 * sizeof(struct Characters));
if (charlist_max) {
    printf("Memory Allocation Error in Charlist_initialize()\n");
    exit(1);
}
charlist_size = 0;
charlist_allocated = charlist_max; /* Size Allocated */
charlist_first = 0; /* No first character */
charlist_vector_height = vector_height;

/* Allocate memory for curves and set pointers */
/* Note: This routine hardcodes the array into the code */
for (i=0; i<charlist_max; i++) {
    charlist->chr[i].curve = i*vector_height;
    charlist->chr[i].curve = i*vector_height;
}

/* .....
 * Zero a previously allocated Characterlist */
void Charlist_zero(struct Charlist *charlist)
{
    struct Charlist *clist;
    charlist_size = 0;
    charlist_first = 0;
}

/* .....
 * Terminate & deallocate the Characterlist */
void Charlist_terminate(charlist)
struct Charlist *charlist;
{
    int i;
    for (i=charlist_allocated-1; i>= 0; i--) { /* Free curves */
        free_vector(charlist->chr[i].curve);
        free_vector(charlist->chr[i].curve);
    }
}

```

```

/* .....
 * Place a new character in the charlist */
/* Takes the character beyond the charlist (charlist),
 * calculates loc, and inserts it into the list */
void Charlist_insert(struct Charlist *charlist,
struct Character *newchar)
/* .....
 * Pointer to the new character adding */
newchar = (struct Character *) malloc(sizeof(struct Character));
/* Calculate the position of the new character */
Charlist_recalculate_loc(newchar);

/* Check for full list */
/* Increase the size of the list */
if (charlist_size == charlist_allocated) /* If too big.. */
    printf("Error in Charlist_place_new_char(): Character list is FULL!\n");
    exit(1);

/* Place new in the list character */
Charlist_locate_character(charlist, newchar);

/* .....
 * This routine repositions a character in the charlist. It removes
 * the character, recalculates loc and inserts it in the list accordingly
void Charlist_relocate_char(charlist, ch)
struct Charlist *charlist;
struct Character *ch;
/* .....
 * Recalculate the position of the character */
Charlist_recalculate_loc(ch);
if (charlist_size < 2) /* If it's the only item in the list */
    return;

/* Remove the character from the list */
if (ch->prev == 0) /* If it's the first item on list */
    charlist->first = charlist->first->next;
else {
    ch->prev->next = ch->next;
    if (ch->next != 0) ch->next->prev = ch->prev;
}

/* Place the character in the list */
Charlist_locate_character(charlist, ch);

/* .....
 * Insert a character after doing a cut */

```

【図 2 2 2】

```

/* This routine will insert the character beyond the charlist (chr[size]).
 * loc is recalculated for both character and the new character, however
 * loc is NOT USED TO PLACE THE CHARACTER. It will
 * place the new character after character after finished, it is an
 * extremely good idea to reset the breakflag!
 */
void Charlist_Insert_After_Put(Charlist, ch, after)
{
    struct Charlist *charlist; /* Pointer to the new character adding */
    struct Character *newchar; /* Set newchar */
    newchar = (struct Character *) malloc(sizeof(struct Character));
    /* Calculate the position of the new character */
    Charlist_Reset_List_Location(charlist, after);
    /* Check for full list */
    if (charlist->size == charlist->alloc) /* If too big... */
    {
        printf("Error in Charlist_Place_New_Char(): Character list is FULL!\n");
        exit(1);
    }
    newchar->prev = charlist->prev;
    newchar->next = charlist->next;
    charlist->next = newchar;
    if (newchar->next != 0)
        newchar->next->prev = newchar;
}

/* This routine will remove a character from the character list
 * Note: This memory is NOT reclaimed!! So don't rely on this too much.
 * Also, prev and next pointers of ch are NOT altered!
 */
void Charlist_Remove_Charlist(Charlist, ch)
{
    struct Character *ch;
    /* Remove the character from the list */
    if (ch->prev == 0) /* If 1st on list */
        charlist->first = charlist->first->next;
    else
        ch->prev->next = ch->next;
    if (ch->next != 0)
        ch->next->prev = ch->prev;
}

/* This routine will recalculate the position of a character */
void Charlist_Reset_List_Location(struct Charlist *charlist, int position)
{
    struct Character *ch;
    int i, loc;
    /* Calculate the position of the new character */
    loc = (ch->top) - (ch->bot); /* Position - average position.. */
    loc += ch->curval[i] - ch->curval[1]; /* .. of the curves */
    loc = loc * 4 / (ch->bot - ch->top); /* Set it */
    ch->loc = loc;
}

```

```

/* This routine will place a character (that is not in the charlist)
 * into the charlist.
 */
void Charlist_Place_Character(Charlist, ch)
{
    struct Character *p; /* The character */
    struct Character *last_p; /* Used for searching for location */
    struct Character *first; /* The character before p */
    int loc;
    /* Is the list empty? */
    if (charlist->size == 1)
    {
        ch->next = 0;
        ch->prev = 0;
        charlist->first = ch;
        return;
    }
    /* Possible optimization: Don't start from start of list...
     * start from where at */
    loc = ch->loc; /* Retrieve position */
    /* Is it at the beginning of the list? */
    p = charlist->first; /* Load p with the first item on the list */
    if (loc < p->loc) /* If it's the first on the list */
    {
        p->prev = ch;
        charlist->first = ch; /* Set the pointers */
        ch->prev = 0;
        return;
    }
    /* Set the 1st on the list to this */
    while (p->loc <= loc) /* Until find record dist where to place */
    {
        last_p = p; /* Remember last position */
        p = p->next; /* Get the next character */
        if (p == 0) /* If it's the end of list, insert it there */
            break;
    }
    ch->prev = last_p; /* Insert new position after last_p */
    last_p->next = ch;
    if (p != 0) /* If nextchar isn't at the end of the list */
        p->prev = ch;
}

/* Allocate memory for list */
/* size = size; /* Compute memory allocated, etc */
/* charlist->list = malloc(size); /* Allocate memory */
/* charlist->size = size; /* Set max allocated */
/* charlist->size = 0; /* Set to zero elements */

```

•

—403—

【図225】

```

// pixel = (unsigned char *)malloc((unsigned int) size*sizeof(unsigned char)); //
// Allocate memory
if (pixel == NULL) // Allocate memory
{
    cerr << "Error: Memory Allocation Error: out of memory\n";
    exit(1);
}
for (int i = 0; i < size, i++)
{
    pixel[i] = (unsigned char *)malloc((unsigned int) size*sizeof(unsigned char));
    // Allocate core mem
    if (pixel[i] == NULL)
    {
        cerr << "Error: Memory Allocation Error: out of memory\n";
        exit(1);
    }
}
allocated_size = size;
allocated_sizeC = sizeC;

// Deallocate memory for an image from memory
void cimage::deallocate()
{
    if ((allocated_size != 0) && (allocated_sizeC != 0))
    {
        for (int i = 0; i < allocated_size, i++)
        {
            delete pixel[i];
            if (sizeC != 0) pixel[i] = 0; // Zero out the sizes
            sizeC = 0;
            allocated_size = 0;
            allocated_sizeC = 0;
        }
    }
}

// This allow allocation of memory from outside the class. It will
// Allocate memory if it's larger. If not, it will just adjust the size
void cimage::reallocate(int R, int C)
{
    int new_R, new_C;
    if ((R > allocated_size) || (C > allocated_sizeC))
    {
        new_R = allocated_size;
        new_C = allocated_sizeC;
        if (R > new_R)
            new_R = R;
        if (C > new_C)
            new_C = C;
        allocate(new_R, new_C); // Re-allocate the image
        sizeR = R;
        sizeC = C;
    }
}

// Create a cimage with image value, value x and y are the size
// void cimage::cimage(char value, int R, int C)

```

```

// Filename: cimage.cpp
// This file contains a vast array image utility functions, for a cimage
//
//
//
// Underline this to make images.c run invisibly (except for error)
// Define PRINT_IMAGES_INFO
#include "cimage.h"
#include "m.h"
//.....
//... CLASS: cimage (in array of char)
//.....
cimage::cimage()
{
    sizeR = 0;
    sizeC = 0;
    allocated_size = 0;
    allocated_sizeC = 0;
    dpi = 0;
}
cimage::cimage(int r, int c)
{
    sizeR = 0;
    sizeC = 0;
    allocated_size = 0;
    allocated_sizeC = 0;
    dpi = 0;
    allocate(r, c);
}
cimage::cimage(cimage img)
{
    allocate(sizeR, sizeC); // Allocate memory for new cimage (loads sizes)
    for (int i = 0; i < sizeR, i++)
    {
        for (int j = 0; j < sizeC, j++) // Copy old image to new cimage
            pixel[i][j] = img.pixel[i][j];
    }
}
// Delete the cimage
cimage::~cimage()
{
    deallocate();
}
// Allocate memory for an image from memory
void cimage::allocate(int r, int c)
{
    deallocate(); // Remove current cimage
    if (r < 0 || c < 0)
    {
        cerr << "Illegal cimage size: " << r << " by " << c << "\n";
        exit(1);
    }
    sizeR = r;
    sizeC = c;
    // Set cimage sizes
    if ((r != 0) && (c != 0))

```

【図226】

```

        for (j=C1; j<C2; j++)
            temp_pixel[(i-1)*j+C1] = pixel[i][j];
        allocate(temp_size, temp_sizeC); // Copy temp image to current image
        for (l=0; l<sizeC; l++)
            temp_image[l] = pixel[i][l];
        pixel[i][j] = temp_pixel[i][j];
    }

    // Read in the TIFF file
    // Read in the TIFF file
    void change_read(tifchar *filename);
    {
        FILE *tagfile; // Input file
        long int tagfile_position; // The our file position (in bytes)
        long int width; // Width of picture
        long int height; // Height of picture
        long int bps; // Bytes of black pixel
        long int bps; // Bytes of black pixel
        long int bps; // Bytes of black pixel
        long int raster_offset; // Start of image data
        long int raster_offset; // Start of image data
        long int raster_offset; // Start of image data
        long int i, j; // Tag values
        if ((tagfile = fopen(filename, "rb")) == NULL)
        {
            cout << "UnError in opening " << filename << "\n";
            exit(1);
        }
        tagfile_position = 0;

        // Read in header and make sure it's Intel
        i = getbytes(tagfile, tagfile_position);
        if (i != 0x0001)
        {
            cout << "... Error: TIFF File not in Intel Format\n";
            exit(1);
        }

        // Read in Magic number
        i = getbytes(tagfile, tagfile_position);
        if (i != 0x002A)
        {
            cout << "BAD Magic Number for TIFF file!\n";
            exit(1);
        }

        // Find First file directory
        j = getbytes(tagfile, tagfile_position);
        while(tagfile_position < j)
        {
            i = getbytes(tagfile, tagfile_position);

            // Tags in Directory
            numtags = getbytes(tagfile, tagfile_position);
            while (i-->numtags > 0) {
                tag = getbytes(tagfile, tagfile_position);
                type = getbytes(tagfile, tagfile_position);
                len = getbytes(tagfile, tagfile_position);
                value = getbytes(tagfile, tagfile_position);
                if ((len % 4) != 0) {tag & 3; tag & 3; tag & 3; tag & 3;}
                cout << "Error: Tag length is " << len << "\n";
                exit(1);
            }
        }
    }
}

```

```

        allocate(R,C);
        for (int i=0; i<sizeR; i++)
            for (int j=0; j<sizeC; j++)
                pixel[i][j] = value;
    }

    // Copy a section of a image into another. image = reference image
    // Range in R1, R2, C1, C2
    void change_clip(imageR image, int R1, int R2, int C1, int C2)
    {
        int r,c,t;
        if (R2 < R1)
            // Swap R1 & R2
            (R1, R2) = (R2, R1);
        if (C2 < C1)
            // Swap C1 & C2
            (C1, C2) = (C2, C1);
        // Check limits
        if (R1 < 0) R1 = 0;
        if (R2 < 0) R2 = 0;
        if (C1 < 0) C1 = 0;
        if (C2 < 0) C2 = 0;
        if (R1 > image-sizeR) R1 = image-sizeR;
        if (R2 > image-sizeR) R2 = image-sizeR;
        if (C1 > image-sizeC) C1 = image-sizeC;
        if (C2 > image-sizeC) C2 = image-sizeC;
        allocate(R2-R1, C2-C1);
        dbl = image.dbl;
        for (r=0; r<R2-R1; r++)
            for (c=0; c<C2-C1; c++)
                pixel[r][c] = image.pixel[r+R1][c+C1];
    }

    // Clip image, reduce the size of an image by removing the blank
    // space around it. This adjusts the image size.
    void change_clip_image()
    {
        int R1, R2, C1, C2; // The starting and stopping rows of new image
        int r,c; // Temporary image
        struct image temp; // Temporary image
        R1 = sizeR;
        R2 = 0;
        C1 = sizeC;
        C2 = 0;
        for (i=0; i<sizeR; i++)
            if (pixel[i][0] != 0)
            {
                if (i < R1) R1 = i;
                if (i > R2) R2 = i;
                if (i < C1) C1 = i;
                if (i > C2) C2 = i;
            }
        // Compute dimensions of new image
        if (R1 > R2) R1 = R2 = 0;
        if (C1 > C2) C1 = C2 = 0;
        temp.allocate(R2-R1, C2-C1);
        for (i=R1; i<R2; i++)
            for (j=C1; j<C2; j++)
                temp_image[i-R1][j-C1] = pixel[i][j];
    }
}

```

【図227】

```

switch (tag)
{
    case 0x00FF: // NewSubfileType
        if (value != 0)
        {
            cout << "Can't Deal With Non-Standard NewSubfileType\n";
            exit(1);
        }
        break;
    case 0x00FF: // NewSubfileType
        if (value != 1)
        {
            cout << "Subfile Type not full resolution\n";
            exit(1);
        }
        break;
    case 0x0000: // Width
        break;
    case 0x0001: // Length
        break;
    case 0x0002: // Bits per sample
        break;
    case 0x0003: // Error. Bits/sample must be 11;
        cout << "Error. Bits/sample must be 11";
        exit(1);
        break;
    case 0x0003: // Compression
        if (value != 1)
        {
            cout << "Compression Used! Error\n";
            exit(1);
        }
        break;
    case 0x0006: // Black is
        break;
    case 0x0007: // Thresholding tag
        if (value < 1) (value > 3)
        {
            cout << "Thresholding Tag = Illegal value\n";
            exit(1);
        }
        break;
    case 0x0011: // Offset to raster data
        raster_offset = value;
        break;
    case 0x0017: // Bits = value;
        break;
    case 0x0018: // Bits = value;
        break;
    case 0x0019: // Bits = value;
        break;
    case 0x001A: // Bits = value;
        break;
    case 0x001B: // Bits = value;
        break;
    case 0x001C: // Bits = value;
        break;
    case 0x001D: // Bits = value;
        break;
    case 0x001E: // Bits = value;
        break;
    case 0x001F: // Bits = value;
        break;
    default:
        if ((value & 32768) == 0)
        {
            cout << "Warning: An unsupported tag appears in the TIF file\n";
        }
        break;
} // switch
} // while
}

while(tagfile_position == raster_offset)
{
    i = getbytes(tagfile_position);
    //----- READ IN FILE -----
    allocate((int) length, (int) width);
    int w;
    for (i = 0; i < length; i++)
    {
        for (j = 0; j < width; j++)
        {
            if ((i+j) == 0)
            {
                v = getc(tagfile);
                tagfile_position++;
                if (v == EOF)
                {
                    cout << "\n Error: Premature End of File (found in tiff file)\n";
                    exit(1);
                }
            }
            if ((v & 128) == Black*128)
            {
                pixel[i][j] = 0;
            }
            else
            {
                pixel[i][j] = 1;
                v = v < 128;
            }
        }
        fclose(tagfile);
    }
    // Read in a byte. File position is updated
    // eoferror = 1 to error on EOF, 0 not
    int cimage; getbytes(FILE *f, long int file_position, int eoferror)
    {
        int ci;
        ci = getc(f);
        if ((ci == EOF) && (eoferror == 1))
        {
            cout << "\n Error: Premature End of File (in getbytes)\n";
            exit(1);
        }
        //file_position++;
        return ci;
    }
    // Read in an integer from a file. (LSB, MSB)
    long int cimage; getbytes(FILE *f, long int fp)
    {
        int ci, c2;
        ci = getc(f);
        c2 = getc(f);
        return ci * 256 + c2;
    }
    // Read in a long integer from a file. (LSB, MSB)
    long int cimage; getbytes(FILE *f, long int fp)
    {
        long int ci, c2;
    }
}

```


【図229】

```

// Filename: cimages.h
// This is the header file for cimages.C
//
//
// =====
//
// #define NOX_DEBUG 1
// #include <stream.h>
// #include <stdio.h>
// #include <stdlib.h>
// #include <math.h>

class cimage; // Forward Reference

// =====
// *** CLASS: cimage
// =====
class cimage
{
public:
    int      sizeR; // Number of rows (first index of array)
    int      sizeC; // Number of cols. (second index of array)
    int      allocated_sizeR; // Amount allocated
    int      allocated_sizeC; // Amount allocated
    int      dpi; // Dots per inch of the image (0 if unknown)
    unsigned char *pixel; // The binary image (2d array)

    cimage(); // Initialise
    ~cimage(); // Destroy
    cimage(int r, int c); // Initialise with a size
    cimage(cimage& temp);

    void reallocate(int R, int C); // Will allocate memory for an image
    void fill(char value, int R, int C); // Fill image with a constant value
    void clip(cimage& image, int R1, int R2, int C1, int C2);
    void clipmask(); // Clips zero space surrounding image
    void readtiff(char *filename); // Read in a tiff image
    void plot(int x, int y); // Plot binary image
    void plotbox(int x, int y); // Plot with a box around it

private:
    void allocate(int r, int c); // Allocate memory
    void deallocate(); // Deallocate memory
    int getbyte(FILE *f, long int& file_position, int& eoferror);
    long int get2bytes(FILE *f, long int& fp);
    long int get4bytes(FILE *f, long int& fp);
};

```

—408—

[illegible]

[illegible]

[illegible]

【図233】

```

/* This program contains common vector and matrix macros */
#include <stdio.h>
#include <stdlib.h>
/* Routine for error during memory allocation */
void allocation_error()
{
    char *s;
    int i;
    printf("Error: %s\n", s);
}

/* Allocate an integer vector of size [0..n-1] */
int *ivector(int n)
{
    int *v;
    v = (int *)malloc(sizeof(int) * n);
    if (!v) allocation_error("ivector");
    return v;
}

/* Free an integer vector v of size [0..n-1] */
void free_ivector(v)
{
    int *v;
    free((int *) v);
}

/* Allocate an float vector of size [0..n-1] */
float *fvector(int n)
{
    float *v;
    v = (float *)malloc(sizeof(float) * n);
    if (!v) allocation_error("fvector");
    return v;
}

/* Free an float vector v of size [0..n-1] */
void free_fvector(v)
{
    float *v;
    free((float *) v);
}

/* Allocate a ranged vector v of size [m..n-1] */
int *rvector_ranged(int m, int n)
{
    int *v;
    v = (int *)malloc(sizeof(int) * (n-m));
    if (!v) allocation_error("rvector_ranged");
    return (v+m);
}

/* Free an integer ranged vector v of size [m..n-1] */
void free_rvector_ranged(v, m)
{
    int *v;
    free((int *) (v-m));
}

/* Allocate an integer matrix of size [0..r-1][0..c-1] */
int **matrix(int r, int c)
{
    int **m;
    m = (int **)malloc(sizeof(int *) * r);
    for (i=0; i<r; i++) {
        m[i] = (int *)malloc(sizeof(int) * c);
        if (!m[i]) allocation_error("matrix");
    }
    return m;
}

/* Free an integer matrix m of size [0..r-1][0..c-1] */
void free_matrix(m, r)
{
    int **m;
    int r;
    for (i=0; i<r; i++)
        free((int *) m[i]);
    free((int **) m);
}

/* Allocate an char matrix of size [0..c-1][0..r-1] */
char **cmatrix(int r, int c)
{
    char **m;
    int r;
    m = (char **)malloc(sizeof(char *) * r);
    for (i=0; i<r; i++) {
        m[i] = (char *)malloc(sizeof(char) * c);
        if (!m[i]) allocation_error("cmatrix");
    }
    return m;
}

/* Free an character matrix m of size [0..r-1][0..c-1] */
void free_cmatrix(m, r)
{
    char **m;
    int r;
    for (i=0; i<r; i++)
        free((char *) m[i]);
}

```

[illegible]

【図237】

```

int end_of_list, // Loop termination flag
int row_div, // The division with the topmost text
int rangel, rangel, // The test range in the adjacent division (row 1-1)
int rangel, rangel, // The test range in the current division (row 1-1)
int r1, r1, c1, c2, // The row and column of line in page
int min_line_height, // The minimum line height
int i, j, b, r, o, // Loop variables
int i, j, b, r, o, // Loop variables

stopcol = vector<NOVIDIONS>;
stopcol[0] = PAGE_CD; // This is overlaid
startcol = iatix(NOVIDIONS, page_size); // This is overlaid
cursor = vector<NOVIDIONS>;
size = vector<NOVIDIONS>;
include = vector<NOVIDIONS>;

min_line_height = (int) ((float) page_size * PAGE_DP2_MIN_BLOCK_SIZE);

////////// Step 1: Define columns and find histograms //////////
// Find range of columns
startcol[0] = PAGE_CD;
for (i=0; i<NOVIDIONS-1; i++)
{
    c = PAGE_CD + (PAGE_CD - (i+1)) / NOVIDIONS;
    startcol[i+1] = c - OVERLAP;
    stopcol[i] = c + OVERLAP;
}
stopcol[NOVIDIONS-1] = PAGE_CD;
// Find histograms
for (i=0; i<NOVIDIONS; i++) // For each division
{
    make_partial_histogram, startcol[i], stopcol[i], histogram;
}

////////// Step 2: Make test area lists //////////
for (i=0; i<NOVIDIONS; i++)
{
    // Note: Zeros are guaranteed on both ends of the histogram.
    // Initialize
    for (j=PAGE_CD-1; j<PAGE_CD; j++)
    {
        if (histogram[i][j] == 0) // If starting text
        {
            startcol[i][j] = j;
        }
        if (histogram[i][j] != 0) // If stopping text
        {
            stopcol[i][j] = j;
            size[i] = j - i;
        }
    }
}

////////// Step 3: Find the lines //////////
for (i=0; i<NOVIDIONS; i++)
{
    cursor[i] = 0; // Set all pointers to the top of the list
    end_of_list = 0;
    while(end_of_list == 0)
    {
        // Step 3a: Find the highest text
        a = 1; // Going to be high div
        for (i=0; i<NOVIDIONS; i++)
        {
            if (cursor[i] < size[i]) // If valid cursor
            {
                b = startcol[i][cursor[i]] + stopcol[i][cursor[i]];
            }
        }
    }
}

```

【图 238】

—414—

【図239】

```

int border_ignore;
int i;

Charlist_initialize(&clist, MAX_CHARS_PER_LINE, MAX_LINE_HEIGHT);
Process_line.fill(0,10,3500); // Allocate the memory for the lines
xs_win_dim(&WINDOW_SIZE_X, &WINDOW_SIZE_Y); // Get the window size

printf("Input tiff filename? /d2/tif/");
strcpy(filename1, "/d2/tif/toshiba.tif");
scanf("%s", filename1);
strcpy(filename, "/d2/tif/");
strcat(filename, filename1);

auto_initialize(); // Initialize Auto-Input routines

printf("Enter courtcut average width? ");
COURTCUT_AVG_WIDTH = auto_input_int();

printf("Enter courtcut delta? ");
COURTCUT_DELTA = auto_input_int();

page.dpi = 300;

border_ignore = 1;
// cout << "Input border width to ignore (in pixels)? ";
// cin >> border_ignore;

printf("Reading in File... ");
page.read(tiff(filename));
printf("Done.\n");

// sepp_init(page); // Initialize page printing routines
// border_ignore = (int) ((float) page.dpi * DPI_3_BORDER_IGNORE);
// printf("Border ignore = %d\n", border_ignore);

PAGE_R1 = border_ignore; // Define page size
PAGE_R2 = page.sizeX - border_ignore;
PAGE_AD = PAGE_R2 - PAGE_R1;
PAGE_C1 = border_ignore;
PAGE_C2 = page.sizeY - border_ignore;
PAGE_CD = PAGE_C2 - PAGE_C1;

// ***** THIS IS JUST FOR TESTING *****
xs_flush();
printf("Enter 0 to continue? ");
i = auto_input_int();

// clear_scale_and_plot_image(page);

printf("Pausing because page is displayed. Input 0 to segment page? ");
border_ignore = auto_input_int();

while(i) {
    segment_page(page, &clist);
    xs_flush();

    printf("Program finished. Enter 0 to run again? ");
    int ttemp;
    ttemp = auto_input_int();
}

Charlist_terminate(&clist);
auto_terminate();
}

```

【図241】

```

/* File: util.c
COMMENTS:
This file contains general utility routines.
*/
ut_reverse(width,height,buf);
ut_getbits(x,p,n);
ut_ascii_bitmap(width,height,buf);
Usage:
Just call the routines.
*/
/* ----- Includes ----- */
#include <stdio.h>
#include <stdlib.h>
/* ----- Globals ----- */
/* ----- Enums ----- */
/*
Routines: ut_reverse(width,height,buf)
Description: Reverse the bits for display format
Return: Nothing
Author: Rick Lee
*/
ut_reverse(width,height,buf)
int width;
int height;
char *buf;
{
    unsigned char val,new_val;
    int x,y;
    for(x = 0; x < (width*3)*height; x++)
    {
        val = buf[x];
        new_val = 0;
        for(y = 7; y >= 0; y--)
        {
            if((ut_getbits(val,x++,1))
            {
                new_val = 0x01 << y | new_val;
            }
            buf[x] = new_val;
        }
    }
}
/*
Routines: ut_getbits(x,p,n)
Description: Check if the bit is turned on or not
Return: Nothing
Author: Rick Lee
*/
ut_getbits(x,p,n)
    unsigned char x; /* The character to slice bits from */
    int p; /* The beginning of the bit number */
    int n; /* Number of bits to get */
{
    return(x >> (p-1-n)) & -1 < 0 << n;
}
/*
Routines: ut_ascii_bitmap(width,height,buf)
Description: Display the bitmap using ascii characters
Return: Nothing
Author: Rick Lee
*/
ut_ascii_bitmap(width,height,buf)
int width;
int height;
char *buf;
{
    short bit_code = 0x0010;
    short bit_code = 0x0001;
    short x,y;
    for(x = 0; x < height; x++)
    {
        for(y = 0; y < width; y++)
        {
            bits = buf[(x*3)+y];
            if(bits & 0x01 << 8)
            {
                bit_result = bits & bit_code;
                if(bit_result > 0)
                {
                    printf("1");
                }
                else
                {
                    printf("0");
                }
                bit_code = bit_code >> 1;
            }
            bit_code = 0x0080;
        }
        printf("\n");
    }
}

```

【图 2 4 2】

[illegible]

【図243】

```

XCopyArea(XtDisplay(canvas), pix, XtWindow(canvas), line_gc, 0, 0,
          width, height, x, y);

/* Now free the bitmap */
XFreePixmap(XtDisplay(canvas), bitmap);

/* Now free the bitmap */
XFreePixmap(XtDisplay(canvas), pix);

/* Now flush the queue to X-server */
XFlush(XtDisplay(canvas));
}

/*
Routine: xs_clr_win()
Description: Clear the screen
Return: Nothing
Author: Rick Lee
*/
xs_clr_win()
{
    XClearArea(XtDisplay(canvas), XtWindow(canvas), 0, 0, 0, 0, TRUE);
}

/*
Routine: xs_draw_box(x, y, width, height)
Description: Draw a box on screen
Return: Nothing
Author: Rick Lee
*/
xs_draw_box(x, y, width, height)
int x;
int y;
int width;
int height;
{
    XDrawRectangle(XtDisplay(canvas), XtWindow(canvas), line_gc,
                  x, y, width, height);
}

/*
Routine: xs_string
Description: Draw a string on the screen
Return: Nothing
Author: Rick Lee
*/
xs_string(x, y, str)
int x;
int y;
char *str;
{
    XDrawString(XtDisplay(canvas), XtWindow(canvas), line_gc,
               x, y, str, strlen(str));
}

/*
Routine: xs_setcolor
Description: Set the current color
Return: Nothing
Author: Rick Lee
*/
xs_setcolor(color)
int color;
{
    gcv.foreground = color;
    XSetForeground(XtDisplay(canvas), line_gc, color);
}

/*
Routine: xs_win_dim
Description: Get the size of the window
Return: Nothing
Author: Rick Lee
*/
xs_win_dim(width, height)
int *width;
int *height;
{
    Arg args[2];
    Dimension win_width, win_height;

    XtSetArg(args[0], XtNwidth, &win_width);
    XtSetArg(args[1], XtNheight, &win_height);
    XtGetValues(canvas, args, 2);

    *width = win_width;
    *height = win_height;
}

/*
xs_set_cmap(w)
Widget w;
{
    int ncolors;
    XColor colors[256];
    Colormap my_cmap;
    int count = 1;

    Display *d;
    int scr;
    Colormap def_colormap;
    XColor color;
    int xi;
    int the_color = 1000;

    d = XtDisplay(w);
    scr = DefaultScreen(d);
    ncolors = DisplayCells(d, scr);
    def_colormap = DefaultColormap(d, scr);
    for (xi = 0; xi < ncolors; xi++)
    {
        colors[xi].pixel = xi;
    }
}

```

-419-

[illegible]

【図247】

```

// Filename: zero_vpp.c
// This routine contains the lines to find blocks that have non-zero vpp
// (level 1 segmentation)
//
// =====
//
#include 'charlist.h'
#include 'images.h'

// Scan a line quickly (every 1/4 pixel) looking for an 'on' pixel
int line_scan_quick( // Returns 1 for pixel found, 0 otherwise
    cimage* line, // The line to process
    int c) // The column to scan
{
    int r;
    for (r=0; r<line->sizeR; r++)
        if (line->pixel(r)(c) & 1)
            return 1;
    return 0;
}

// Scan a line completely looking for an 'on' pixel
int line_scan_complete( // Returns 1 for pixel found, 0 for otherwise
    cimage* line, // The line
    int c) // The column to scan
{
    int r;
    for (r=0; r<line->sizeR; r++)
        if (line->pixel(r)(c) & 1)
            return 1;
    return 0;
}

// The main routine. Call this to make the list
void extract_zero_vpp_list(
    cimage* line, // The image of the line to segment
    struct NonZeroVppList *list) // A pointer to the list
{
    int c;
    int left, right;
    for (c=0; c<line->sizeC; c++) // Scan the image
    {
        if (line_scan_quick(line, c)) // if found something
        {
            left = right = c; // Look to the left for the start
            while (--c >= 0)
                if (line_scan_complete(line, c) == 0)
                    break;
            left = c+1;
            c = right; // Keep looking to the right
            while (c < line->sizeC)
                if (line_scan_complete(line, c) == 0)
                    break;
            right = c;
            zero_vpp_add(list, left, right); // Add in the new break
        }
    }
}

```

フロントページの続き

(72)発明者 シン・ヤン ワング
 アメリカ合衆国 カリフォルニア州
 92626, コスタメサ ブルマン ストリー
 ト 3188 キヤノン インフォメーション
 システムズ インク. 内

(72)発明者 メザードゥ アール. バエズィー
 アメリカ合衆国 カリフォルニア州
 92626, コスタメサ ブルマン ストリー
 ト 3188 キヤノン インフォメーション
 システムズ インク. 内

(72)発明者 クリストファー エー. シェリック
 アメリカ合衆国 カリフォルニア州
 92626, コスタメサ ブルマン ストリー
 ト 3188 キヤノン インフォメーション
 システムズ インク. 内